ROBUST DETECTION OF FREQUENCY SHIFT KEYING MODULATED SIGNALS

by

Adam Dennis Hall

B.S. Physics, Brigham Young University, 2009

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Electrical Engineering

2013

This thesis for the Master of Science degree by

Adam Dennis Hall

has been approved for the

Electrical Engineering Program

by

Titsa Papantoni, Chair

Miloje (Mike) Radenkovic

Jan Bialasiewicz

March 14, 2013

Hall, Adam Dennis (M.S., Electrical Engineering)

Robust Detection of Frequency Shift Keying Modulated Signals

Thesis directed by Professor Titsa Papantoni

ABSTRACT

We derive and examine the Neyman-Pearson Detection of an FSK modulated waveform in white Gaussian noise. We specifically examine three cases: 1) detection when the waveform is known exactly; 2) detection when the waveform is known down to an unknown initial phase and; 3) detection of the waveform when the energy level is known but the signal is not. We then derive robust detection schemes where the noise comes from a broader class of stochastic processes and includes the possible occurrence of outlier data. We compare these robust detection schemes to their non-robust equivalents under various noise models.

The form and content of this abstract are approved. I recommend its publication.

Approved: Titsa Papantoni

ACKNOWLEDGMENTS

I wish to give my sincere thanks to my advisor, Dr. Titsa Papantoni. She taught me detection and estimation theory which forms the foundation for much of this thesis. In addition, she provided guidance at all stages of writing this thesis from first shaping this work to publishing. And most importantly, she brought out and expanded my passion for learning. I would like to thank Dr. Radenkovic as well as Dr. Bialasiewicz, for their time and effort as part of the committee charged with the approval of this thesis. Finally, I would like to thank each and every one of my professors for sharing their domain knowledge.

DEDICATION

This thesis is dedicated to my wife Rachel, for her support, patience and sacrifice as I pursued this Masters degree and to my son Max, whose smile and laughter brought me joy and strength as I worked on this thesis. Finally, I want to thank my employer, Northrop Grumman Information Systems, for allowing me to pursue this degree and their financial support. After the time and effort dedicated to the completion of this degree and this thesis, I am truly a better engineer and more importantly, I have a deeper understanding of the world around me.

TABLE OF CONTENTS

Chapter

1 Introduction 1
  1.1 Overview 1
  1.2 Notation and Symbols 2
2 Parametric Detection 4
  2.1 Introduction 4
  2.2 Detection when the waveform and initial phase are known 4
  2.3 Detection when the waveform is unknown (energy detection) 6
  2.4 Detection when the waveform is known down to an unknown initial phase 7
  2.5 Comparison of the Detection Performance 15
    2.5.1 Detection Performance of Neyman Pearson-like Detectors 16
    2.5.2 Detection Performance of the Neyman-Pearson Detectors 19
3 Robust Detection 24
  3.1 Detection when the waveform and the initial phase are known 24
    3.1.1 Case where contamination occurs over data pairs 24
    3.1.2 Case where contamination occurs per datum 30
    3.1.3 Comparison of the robust cases with the optimal at the nominal model test 39
  3.2 Detection when the waveform is unknown (energy detection) 50
  3.3 Detection when the waveform is known down to an unknown initial phase 69
    3.3.1 Robust Estimation of the unknown initial phase 70

3.3.2 Robust Detection 78

3.3.3 Comparing robust detection with non-robust detection 83

4 Thesis Conclusions 88

4.1 Summary of Findings 88

4.2 Future Research 90

Bibliography 92

Appendix

A: Extension of Huber Robust Detection 93

1 Introduction

1.1 Overview

Signal Detection is a key step in many communication, RADAR, and signal geolocation systems. Optimal signal detection design is dependent on the apriori knowledge of the signal being detected as well as the correct model for the transmission or noise channel. For any real world application it is extremely difficult to precisely model the noise channel. Furthermore, optimal signal detectors' overall performance can be extremely sensitive to the noise channel. In contrast, robust detectors provide good performance at the assumed noise channel and maintain good performance over a wide class of possible noise channels. This thesis examines non-robust and robust detection schemes for three cases of apriori signal knowledge and provides detailed analysis comparing these detection schemes.

In Chapter 2 we derive and compare optimal signal detection of a Frequency Shift Keying (FSK) modulated signal in an additive white Gaussian noise channel for three typical cases of apriori signal knowledge. The relative performance gain between the cases is found to be significantly different than approximate performance gain figures derived using first and second order statistics only.

In Chapter 3 we derive robust signal detection schemes for the three cases examined in Chapter 2. For two of the three cases, the robust detection schemes derived in Chapter 3 can be shown to be optimal; they are Neyman Pearson detection schemes designed at the least favorable density pair. For the remaining case, an ad-hoc but intuitively pleasing detection scheme is designed and shown to provide good performance in the presence of low frequency outliers. All cases are compared to the non-robust detection schemes found in Chapter 2 under various noise models.

1.2 Notation and Symbols

Throughout this thesis all symbols used to represent variables, random variables, constants, or functions are defined when first introduced. The table below contains a brief description of terms which are common among multiple chapters and sections of this thesis. The table is not an exhaustive list, as many symbols used for a given section are unique to that section and are therefore best introduced when they are first used. Furthermore, the table does not contain a precise definition of all terms listed as such a precise definition can only be provided in the context of the equations in which the terms are used. The table is placed here to introduce the reader to common symbols and enable the reader to understand a given chapter or section without reading all preceding text.

Table 1: Table of mathematical symbols.

E	Energy of the signal per sample
f_i	Probability Density function assuming the i^{th} hypothesis is active
F_i	Cumulative Density function assuming the i^{th} hypothesis is active
$\mathfrak{I}_i$	The class of all possible density functions active if the i^{th} hypothesis is active
H_i	The i^{th} hypothesis
$I_0(x)$	The zeroth ordered Bessel function evaluated at x
n	The number of samples
N_o	The noise spectral density
P_{fa}	Probability of false alarm
P_d	The power or probability of detection

Table 1 (con't.)

$Q_n(a,b)$	The Marcum-Q function of n^{th} order evaluated at (a,b)
x_i	The real part of a complex sample from the observation vector
y_i	The imaginary part of a complex sample from the observation vector
α	False alarm constraint
γ	Threshold value used in detection tests
δ	The decision rule for a given detection test
ε_i	The frequency of outliers assuming the i^{th} hypothesis is active
θ_o	The initial phase offset of the signal
μ	The mean of a random variable
ρ	The signal-to-noise ratio
σ	The standard deviation of a random variable generally Gaussian
υ	The standard deviation of a random variable
ϕ(x)	The probability density function of a zero mean, unit variance Gaussian variable evaluated at x
Φ(x)	The cumulative density function of a zero mean, unit variance Gaussian variable evaluated at x

2 Parametric Detection

2.1 Introduction

After the signal is brought to baseband, it is low-pass filtered and sampled. We then obtain the following two data sample vectors

$$X(i) = A(i) + N_1(i) \qquad \text{(Eq. 2.1.1)}$$

$$Y(i) = B(i) + N_2(i) \qquad \text{(Eq. 2.1.2)}$$

$N_1(i)$ and $N_2(i)$ are Independent White Gaussian Noise processes with $\sigma^2 = N_0$. The expressions for$A(i), B(i)$ depend on the low-pass filter used as shown in [1]. To simplify our presentation, we assume the deployment of an ideal low-pass filter without distortion such that

$$A(i) = \sqrt{E} * \cos\,(\boldsymbol{\theta_o} + P(i)) \qquad \text{(Eq. 2.1.3)}$$

$$B(i) = \sqrt{E} * \sin\left(\boldsymbol{\theta_o} + P(i)\right) \qquad \text{(Eq. 2.1.4)}$$

where $\boldsymbol{\theta_o}$ is a uniform random variable on the interval $[-\pi, \pi]$; for scenarios where it is unknown and a fixed constant for scenarios where it is known. $P(i)$ is in general a random variable; for the case where the waveform is unknown and a deterministic function; for the case where the waveform is known. Finally, E is the energy of the signal.

2.2 Detection when the waveform and initial phase are known

In this case, $P(i)$ is no longer random; equivalently, $Prob\{P(i) = p(i)\} = 1$, for some deterministic $p(i)$. The optimal detector is given by deciding H_1 if

$$\frac{1}{n}\sum_{i=1}^{n} x(i)\cos\left(\theta_o + p(i)\right) + \frac{1}{n}\sum_{i=1}^{n} y(i)\sin\left(\theta_o + p(i)\right) > \gamma. \qquad \text{(Eq. 2.2.1)}$$

Let

$$Z = \frac{1}{n}\sum_{i=1}^{n} x(i)\cos\left(\theta_o + p(i)\right) + \frac{1}{n}\sum_{i=1}^{n} y(i)\sin\left(\theta_o + p(i)\right) \qquad \text{(Eq. 2.2.2)}$$

The detection statistics are then,

$$f(Z|H_1) = \phi\left(\frac{Z - \sqrt{E}}{\sqrt{N_o/n}}\right) \qquad \text{(Eq. 2.2.3)}$$

$$f(Z|H_o) = \phi\left(\frac{Z}{\sqrt{N_o/n}}\right) \qquad \text{(Eq. 2.2.4)}$$

$$P_{fa} = Prob(Z > \gamma|H_o) = 1 - \Phi\left(\frac{\gamma\sqrt{n}}{\sqrt{N_o}}\right) = \Phi\left(-\frac{\gamma\sqrt{n}}{\sqrt{N_o}}\right) = \alpha \qquad \text{(Eq. 2.2.5)}$$

$$\gamma = -\sqrt{N_o/n}\,\Phi^{-1}(\alpha) \qquad \text{(Eq. 2.2.6)}$$

$$P_d = Prob(Z > \gamma|H_1) = 1 - \Phi\left(\frac{\gamma\sqrt{n} - \sqrt{nE}}{\sqrt{N_o}}\right) = \Phi\left(\frac{\sqrt{nE} - \gamma\sqrt{n}}{\sqrt{N_o}}\right) \qquad \text{(Eq. 2.2.7)}$$

$$P_d = \Phi\left(\frac{\sqrt{nE}}{\sqrt{N_o}} + \Phi^{-1}(\alpha)\right) \qquad \text{(Eq. 2.2.8)}$$

Define $\rho = \frac{E}{2N_o}$. And re-express the probability of correct signal detection as follows:

$$P_d = \Phi\left(\sqrt{2n\rho} + \Phi^{-1}(\alpha)\right) \qquad \text{(Eq. 2.2.9)}$$

2.3 Detection when the waveform is unknown (energy detection)

$P(i)$ is now a random variable, which in general depends on the FSK index. If we do not know the FSK index, we will work under the model that each $P(i)$ is a random variable uniformly distributed on $[-\pi, \pi]$. Our optimal detector is now a simpler energy detector given by deciding H_1 if

$$\frac{1}{n}\sum_{i=1}^{n}[x_i^2 + y_i^2] > \gamma \qquad \text{(Eq. 2.3.1)}$$

Our detection statistics are

$$f(Z|H_0) = \frac{n}{2^n(n-1)!\,N_o}\left(\frac{Zn}{N_o}\right)^{n-1} \exp\left(-\frac{Zn}{2N_o}\right) \qquad \text{(Eq. 2.3.2)}$$

$$f(Z|H_1) = \frac{n}{2\mathrm{N_o}} \exp\left(\frac{-Zn - nE}{2N_o}\right)\left(\frac{Z}{E}\right)^{\frac{n}{2}-\frac{1}{2}} I_{n-1}\left(\frac{\sqrt{n^2EZ}}{N_o}\right) \qquad \text{(Eq. 2.3.3)}$$

$$P_{fa} = Prob(Z > \gamma|H_0) = \int_{\gamma}^{\infty} dZ \frac{n}{2^n(n-1)!\,N_o}\left(\frac{nZ}{N_o}\right)^{n-1} \exp\left(-\frac{nZ}{2N_o}\right) = \alpha \qquad \text{(Eq. 2.3.4)}$$

Let $x = Zn/N_o$

$$P_{fa} = \int_{\gamma n/N_o}^{\infty} dx \frac{1}{2^n(n-1)!}(x)^{n-1} \exp\left(-\frac{x}{2}\right) = \alpha \qquad \text{(Eq. 2.3.5)}$$

Define

$$F_n(x) = \int_{0}^{x} d\tau \frac{1}{2^n(n-1)!}(\tau)^{n-1} \exp\left(-\frac{x}{2}\right) \qquad \text{(Eq. 2.3.6)}$$

$F_n(x)$ is the cumulative distribution function of the gamma random variable with parameters 2 and n.

$$P_{fa} = 1 - F_n\left(\frac{n\gamma}{N_o}\right) = \alpha \qquad \text{(Eq. 2.3.7)}$$

$$\gamma = \frac{N_o}{n} F_n^{-1}(1-\alpha) \qquad \text{(Eq. 2.3.8)}$$

$$P_d = Prob(Z > \gamma | H_1) = Q_n\left(\sqrt{\frac{nE}{N_o}}, \sqrt{\frac{n\gamma}{N_o}}\right) \qquad \text{(Eq. 2.3.9)}$$

Let $\rho = \frac{E}{2N_o}$. Then,

$$P_d = Q_n\left(\sqrt{2n\rho}, \sqrt{F_n^{-1}(1-\alpha)}\right) \qquad \text{(Eq. 2.3.10)}$$

2.4 Detection when the waveform is known down to an unknown initial phase

In this case $P(i)$ is no longer random. Equivalently, there exists a deterministic $p(i)$ such that $Prob\{P(i) = p(i)\} = 1$. We will design the Neyman-Pearson Detector for such a model. The Neyman-Pearson Detector decides H_1 if

$$\frac{f_{H1}(x^n, y^n)}{f_{H0}(x^n, y^n)} \geq \gamma \qquad \text{(Eq. 2.4.1)}$$

while H_0 is decided otherwise, where x^n, y^n are the observed data from the processes $X(n), Y(n)$, where $f_{Hi}(x^n, y^n)$, is the probability density functions when H_i is active, and where γ is a threshold determined by some false alarm constraint α. We will show that $f_{Hi}(x^n, y^n)$ does exist for both $i = 0,1$. To find $f_{H1}(x^n, y^n)$, we note by the theorem of total probability:

$$f_{H1}(x^n, y^n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{H1}(x^n, y^n | \boldsymbol{\theta_o} = \theta) d\theta \qquad \text{(Eq. 2.4.2)}$$

And

$$f_{H1}(x^n, y^n|\boldsymbol{\theta_o} = \theta) = \prod_{i=1}^{n} f_{H1}(x_i, y_i|\boldsymbol{\theta_o} = \theta) = \prod_{i=1}^{n} f_{H1}(x_i|\boldsymbol{\theta_o} = \theta)\, f_{H1}(y_i|\boldsymbol{\theta_o}$$

(Eq. 2.4.3)

$$= \theta) = \prod_{i=1}^{n} \phi(\frac{x_i - \sqrt{E} * \cos(\theta + p(i))}{\sqrt{N_0}})\ \phi(\frac{y_i - \sqrt{E} * \sin(\theta + p(i))}{\sqrt{N_0}})$$

where $\phi(x)$ is the probability density function for the zero-mean unit Gaussian random variable.

$$f_{H1}(x^n, y^n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{H1}(x^n, y^n|\, \boldsymbol{\theta_o} = \theta) d\theta =$$

(Eq. 2.4.4)

$$\frac{1}{2\pi}\frac{1}{N_o^n}\frac{1}{(2\pi)^n}\int_{-\pi}^{\pi} \exp\,(-\frac{1}{2N_0}\sum_{i=1}^{n}(x_i - \sqrt{E}\cos(\theta + p(i)))^\wedge 2 +$$

(Eq. 2.4.5)

$$(y_i - \sqrt{E}\sin\big(\theta + p(i)\big))^\wedge 2\,)\; d\theta =$$

$$c * \exp\,(-\frac{1}{2N_0}(\sum_{i=1}^{n}[x_i^2 + y_i^2] + nE)\,) *$$

(Eq. 2.4.6)

$$\int_{-\pi}^{\pi} \exp\,(\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}[x_i\ \cos(\theta + p(i))\ + y_i\ \sin(\theta + p(i))]\,)d\theta$$

where we have defined, $c = \frac{1}{N_o^n}\frac{1}{(2\pi)^{n+1}}$. In contrast to the detection case where both the waveform and the phase offset are known, in this case, where the waveform is known but not the phase offset, $f_{H1}(x^n, y^n)$ is not memoryless, since

$$\begin{aligned} f_{H1}(x^n, y^n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{H1}(x^n, y^n \mid \boldsymbol{\theta_o} = \theta)\,d\theta \neq \\ f_{H1}(x_n\ \ , y_n)&\frac{1}{2\pi}\int_{-\pi}^{\pi} f_{H1}(x^{n-1}, y^{n-1} \mid \boldsymbol{\theta_o} = \theta)\,d\theta \\ &= f_{H1}(x_n\ \ , y_n) f_{H1}(x^{n-1}, y^{n-1}\ \ ) \end{aligned} \tag{Eq. 2.4.7}$$

To ultimately simplify the detection scheme, we now seek to simplify the expression for $f_{H1}(x^n, y^n)$. We will start with the integrand

$$\exp\,(\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}[x_i\ \cos(\,\theta + p(i))\ + y_i\ \sin(\,\theta + p(i))]\,) \tag{Eq. 2.4.8}$$

Let $\mu_i = \tan^{-1}\left(\frac{y_i}{x_i}\right)$. The integrand above can be rewritten as

$$\exp(\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[\left[x_i^2 + y_i^2\right]^{\frac{1}{2}}\{\cos\,(\mu_i)\ \cos(\,\theta + p(i))\ + \sin\,(\mu_i)\ \sin(\,\theta + p(i))\}\,\right]) = \tag{Eq. 2.4.9}$$

$$\exp\left(\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[x_i^2 + y_i^2\right]^{\frac{1}{2}}\cos\,(p(i) + \theta - \mu_i)\ \right) = \tag{Eq. 2.4.10}$$

$$\exp\left(\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[x_i^2 + y_i^2\right]^{\frac{1}{2}} Re\{\exp\,(\mathrm{j}p(i) + \mathrm{j}\theta - \mathrm{j}\mu_i)\}\ \right) = \tag{Eq. 2.4.11}$$

$$\exp\left(Re\{\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[x_i^2 + y_i^2\right]^{\frac{1}{2}}\exp\,(\mathrm{j}p(i) + \mathrm{j}\theta - \mathrm{j}\mu_i)\}\ \right) = \tag{Eq. 2.4.12}$$

$$\exp\left(Re\{\exp\,(j\theta)\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[x_i^2 + y_i^2\right]^{\frac{1}{2}}\exp\,(\mathrm{j}p(i) - \mathrm{j}\mu_i)\}\ \right) \tag{Eq. 2.4.13}$$

where $Re\{x\}$ denotes the real part. Let

$$h(i) = \exp\,(jp(i)) \qquad \text{(Eq. 2.4.14)}$$

$$T_i = \left[x_i^2 + y_i^2\right]^{\frac{1}{2}} \exp\,(j\mu_i) \qquad \text{(Eq. 2.4.15)}$$

$$T = |\sum_{i=1}^{n} T_i\,| \qquad \text{(Eq. 2.4.16)}$$

$$V = \left|\sum_{i=1}^{n} T_i \quad h^*(i)\right| \qquad \text{(Eq. 2.4.17)}$$

where s^* denotes the conjugate of s. Finally, define ω such that

$$\sum_{i=1}^{n} T_i \quad h^*(i) = Ve^{-j\omega} \qquad \text{(Eq. 2.4.18)}$$

With these substitutions we have

$$\frac{\sqrt{E}}{N_0}\sum_{i=1}^{n}\left[x_i^2 + y_i^2\right]^{\frac{1}{2}} \exp\,(\mathrm{j}p(i) - \mathrm{j}\mu_i) \; = \frac{\sqrt{E}}{N_0} Ve^{j\omega} \qquad \text{(Eq. 2.4.19)}$$

The overall integral we wish to evaluate is

$$f_{H1}(x^n, y^n) = c * \exp\,(-\frac{1}{2N_0}[T^2 + nE]) \int_{-\pi}^{\pi} \exp\,(\frac{\sqrt{E}}{N_0} V * Re\{\,\exp(j[\omega + \theta])\}\,) d\theta \qquad \text{(Eq. 2.4.20)}$$

Finally, with the change of variables $= \omega + \theta$, we write,

$$c * \exp\,(-\frac{1}{2N_0}[T^2 + nE])\int_{-\pi+\theta}^{\pi+\theta} \exp\,(\frac{\sqrt{E}}{N_0}V * Re\{\exp(j\tau)\}\,)d\tau = \qquad \text{(Eq. 2.4.21)}$$

$$c * \exp\,(-\frac{1}{2N_0}[T^2 + nE])\int_{-\pi}^{\pi} \exp\,(\frac{\sqrt{E}}{N_0}V * \cos\,(\tau))d\tau = \qquad \text{(Eq. 2.4.22)}$$

$$2c * \exp\,(-\frac{1}{2N_0}[T^2 + nE])\int_{0}^{\pi} \exp\,(\frac{\sqrt{E}}{N_0}V * \cos\,(\tau))d\tau \qquad \text{(Eq. 2.4.23)}$$

Using the identity

$$\frac{1}{\pi}\int_{0}^{\pi} \exp\,(r\cos\,(\tau))d\tau = I_o(r) \qquad \text{(Eq. 2.4.24)}$$

We rewrite (Eq. 2.4.23) as,

$$f_{H1}(x^n, y^n) = c'\exp\,(-\frac{1}{2N_0}[T^2 + nE])I_o(\frac{\sqrt{E}}{N_0}P) \qquad \text{(Eq. 2.4.25)}$$

where $c' = 2\pi c = \frac{1}{N_o^n}\frac{1}{(2\pi)^n}$. The likelihood ratio is thus

$$\exp\left(-\frac{nE}{2N_0}\right)I_o\left(\frac{\sqrt{E}}{N_0}V\right) \qquad \text{(Eq. 2.4.26)}$$

Taking the log ratio and arranging terms, our test decides H_1 if

$$\log\left[I_o\left(\frac{\sqrt{E}}{N_0}V\right)\right] \geq \gamma' \qquad \text{(Eq. 2.4.27)}$$

So far our test takes the observed sequence $T_i = x_i + jy_i$, and the known transmitted signal (with unit amplitude) $h(i)$, and computes

$$V = \left| \sum_{i=1}^{n} T_i \quad h^*(i) \right| \qquad \text{(Eq. 2.4.28)}$$

Then, our test computes $\log\left[I_o\left(\frac{\sqrt{E}}{N_0}V\right)\right]$, and compares this test statistic to some threshold γ' selected to attain probability of false alarm equal to α. However, since the function $\log\left[I_o\left(\frac{\sqrt{E}}{N_0}V\right)\right]$ is strictly a function of V and is monotonically increasing with V, for $V > 0$, while V is by definition always nonnegative, the detection rule

$$\log\left[I_o\left(\frac{\sqrt{E}}{N_0}V\right)\right] \geq \gamma(\alpha) \qquad \text{(Eq. 2.4.29)}$$

reduces to the equivalent inequality,

$$V \geq \gamma''(\alpha) \qquad \text{(Eq. 2.4.30)}$$

In other words, V is the sufficient statistics of the detection scheme. Finally, we can multiply both sides of the detection inequity by $\frac{1}{n}V$ (we do this to more easily derived the analytic expression for the probability of detection). In conclusion, our detection rule decides H_1 if

$$\frac{1}{n}V^2 \geq \gamma''(\alpha) \qquad \text{(Eq. 2.4.31)}$$

where

$$V = \left| \sum_{i=1}^{n} T_i \quad h^*(i) \right| \qquad \text{(Eq. 2.4.32)}$$

and

$$T_i = x_i + jy_i \qquad \text{(Eq. 2.4.33)}$$

$$h(i) = \exp\big(jp(i)\big) \qquad \text{(Eq. 2.4.34)}$$

A similar derivation for the Bayesian Detection of several orthogonal FSK signals is found in [2]. To evaluate the performance of our detection scheme, we now seek an analytic expression for the probability density function of $\frac{1}{n}\boldsymbol{V}^2$ under both hypotheses. Define $\boldsymbol{Z} = \frac{1}{n}\boldsymbol{V}^2$ to be the random variable we seek to find under both H_0 and H_1. We first examine the statistics of

$$\sum_{i=1}^{n} \boldsymbol{T}_i \;\; h^*(i) \qquad \text{(Eq. 2.4.35)}$$

which is a linear combination of complex, independent Gaussian variables. Since any linear combination of Gaussian variables is Gaussian, the

$$\sum_{i=1}^{n} \boldsymbol{T}_i \;\; h^*(i) \qquad \text{(Eq. 2.4.36)}$$

is a complex Gaussian variable and can be completely described by its mean and variance

$$E\left\{\sum_{i=1}^{n} \boldsymbol{T}_i \;\; h^*(i)\right\} = \mathrm{n}\sqrt{\mathrm{E}}(\cos\theta_o + j\sin\theta_o) \qquad (\text{under } H_1) \qquad \text{(Eq. 2.4.37)}$$

$$E\left\{\sum_{i=1}^{n} \boldsymbol{T}_i \;\; h^*(i)\right\} = 0 \qquad (\text{under } H_o) \qquad \text{(Eq. 2.4.38)}$$

$$Var\left\{\sum_{i=1}^{n} \boldsymbol{T}_i \;\; h^*(i)\right\} = nN_0 + jnN_0 \qquad (\text{under both } H_1 \text{ and } H_o) \qquad \text{(Eq. 2.4.39)}$$

Therefore under $H_{0,}$ the distribution of $\boldsymbol{Z}$ is $chi - square$, and under H_1, the distribution of $\boldsymbol{Z}$ is generalized $chi - square$. Specifically

$$f(Z|H_o) = \frac{1}{2N_o}\exp\left(-\frac{Z}{2N_o}\right) \qquad \text{(Eq. 2.4.40)}$$

$$f(\mathrm{Z}|\mathrm{H}_1) = \frac{1}{2\mathrm{N_o}}\exp\left(\frac{-Z - nE}{2N_o}\right) I_o\left(\frac{\sqrt{nEZ}}{N_o}\right) \qquad \text{(Eq. 2.4.41)}$$

$$Probability\ False\ alarm = Prob\{\ \boldsymbol{Z} > \gamma|H_o\} = 1 - \int_0^{\gamma} f(Z|H_o)\, dZ = \qquad \text{(Eq. 2.4.42)}$$

$$1 - \left(1 - \exp\left(-\frac{\gamma}{2N_o}\right)\right) = \exp\left(-\frac{\gamma}{2N_o}\right) = \alpha \qquad \text{(Eq. 2.4.43)}$$

$$Probability\ Detection = Prob\{\ \boldsymbol{Z} > \gamma|H_1\} = 1 - \int_0^{\gamma} f(Z|H_1)\, dZ = \qquad \text{(Eq. 2.4.44)}$$

$$\int_y^{\infty} \frac{1}{2\mathrm{N_o}}\exp\left(\frac{-Z + nE}{2N_o}\right) I_o\left(\frac{\sqrt{nEZ}}{N_o}\right) dZ \qquad \text{(Eq. 2.4.45)}$$

Let $x = \sqrt{Z/N_o}$

$$\int_{\sqrt{\frac{\gamma}{N_o}}}^{\infty} \mathrm{x}\exp\left(-\frac{x^2}{2} - \frac{nE}{2N_o}\right) I_o\left(x\sqrt{\frac{nE}{N_o}}\right) dx = Q_1\left(\sqrt{\frac{nE}{N_o}}, \sqrt{\frac{\gamma}{N_o}}\right) \qquad \text{(Eq. 2.4.46)}$$

where $Q_1(a, b)$ is the Marcum-Q-function with $M = 1$. Define ρ to be the signal power to noise power ratio, per complex sample, where $\rho = \frac{E}{2N_o}$. The probability of detection can be written as a function of the number of samples, the signal to noise ratio ρ and the probability of false alarm α, as:

$$P_d = Q_1\left(\sqrt{2n\rho}, \sqrt{2\ln(1/\alpha)}\right) \qquad \text{(Eq. 2.4.47)}$$

Figure 2-1 shows the probability of detection versus the product np for various false alarms.

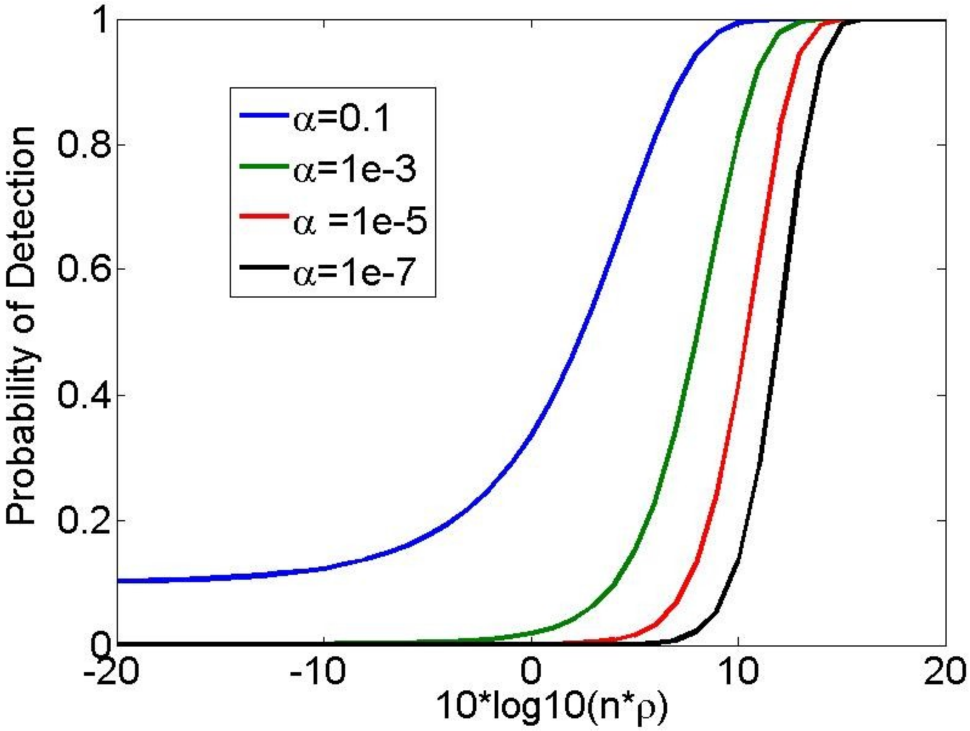

Figure 2-1: Power curves as a function of $n\rho$ for varies false alarm rates α.

2.5 Comparison of the Detection Performance

In this section, we will compare the detection performance for all three cases examined in sections 2.2, 2.3, and 2.4, by examining the induced probabilities of detection, as functions of α,ρ, and n. However, we will first compare the detection performance of Neyman-Pearson-like detectors which use just the mean and variance of our test statistic under H_0 and H_1, for all three cases. This analysis, being based on first and second order statistics, will lead to simpler expressions regarding comparisons between the three cases; however, the latter expressions are simply wrong when used for evaluation of the Optimal Neyman-Pearson detectors. The reason we present the performance of detectors that use just the mean and the variance is to demonstrate the danger involved in limiting performance evaluations to just first and second order statistics.

2.5.1 Detection Performance of Neyman Pearson-like Detectors

Let μ_{Hi} and σ_{Hi} be the mean and standard deviation of our test statistic under the acting hypothesis H_i, where i can be either 0 or 1. Let us design our detection threshold to be $\gamma = \mu_{H0} + m\sigma_{H0}$, where m is a constant. In other words, we set our threshold to be $m\sigma$ away from the mean, μ, under H_o. Clearly, the greater the selected m, the lower the resulting probability of false alarm. Furthermore, to simplify our detection performance analysis we will examine the values of n, and ρ needed for the case $\mu_{H1} = \gamma$ in each of the three cases. We are thus using the first and second order statistics of our test statistic under H_0 and just the mean under H_1 to compare the three cases.

From section 2.2, in the case where the waveform as well as the initial phase are known, we have

$$\mu_{H0} = 0 \qquad \text{(Eq. 2.5.1)}$$

$$\sigma_{H0} = \sqrt{N_o/n} \qquad \text{(Eq. 2.5.2)}$$

$$\mu_{H1} = \sqrt{E} \qquad \text{(Eq. 2.5.3)}$$

Our detection comparison occurs when $\mu_{H1} = \mu_{H0} + m\sigma_{H0}$ or

$$\frac{nE}{N_o} = m \qquad \text{(Eq. 2.5.4)}$$

or equivalently when

$$2n\rho = m. \qquad \text{(Eq. 2.5.5)}$$

From section 2.3 Detection when the waveform is unknown (energy detection), we have

$$\mu_{H0} = 2N_o \qquad \text{(Eq. 2.5.6)}$$

$$\sigma_{H0} = \frac{2N_o}{\sqrt{n}} \qquad \text{(Eq. 2.5.7)}$$

$$\mu_{H1} = 2N_o + E \qquad \text{(Eq. 2.5.8)}$$

Our detection comparison occurs when $\mu_{H1} = \mu_{H0} + m\sigma_{H0}$ or

$$\frac{E}{2N_o}\sqrt{n} = m \qquad \text{(Eq. 2.5.9)}$$

or equivalently when

$$\rho\sqrt{n} = m \qquad \text{(Eq. 2.5.10)}$$

From section 2.4, Detection when the waveform is known down to an unknown initial phase, we have

$$\mu_{H0} = 2N_o \qquad \text{(Eq. 2.5.11)}$$

$$\sigma_{H0} = 2N_o \qquad \text{(Eq. 2.5.12)}$$

$$\mu_{H1} = 2N_o + nE \qquad \text{(Eq. 2.5.13)}$$

Our detection comparison occurs when $\mu_{H1} = \mu_{H0} + m\sigma_{H0}$ or

$$\rho n = m \qquad \text{(Eq. 2.5.14)}$$

To compare the three cases, let $\omega_{coh}(p, n)$, $\omega_{energy}(p, n)$ and $\omega_{incoh}(p, n)$ equal the expressions on the left hand side of (Eq. 2.5.5), (Eq. 2.5.10) and (Eq. 2.5.14), respectively. We then obtain the following ratios, as comparison quantities between the three tests

$$\frac{\omega_{coh}(p, n)}{\omega_{energy}(p, n)} = 2\sqrt{n} \qquad \text{(Eq. 2.5.15)}$$

$$\frac{\omega_{coh}(p, n)}{\omega_{incoh}(p, n)} = 2 \qquad \text{(Eq. 2.5.16)}$$

$$\frac{\omega_{incoh}(p, n)}{\omega_{energy}(p, n)} = \sqrt{n} \qquad \text{(Eq. 2.5.17)}$$

Based on first and second order statistics, our analysis results in simple and easy-to-remember detection comparisons. From (Eq. 2.5.16) we conclude that the coherent detector performs approximately 3 dB better than the incoherent detector (in either ρ or n). From (Eq. 2.5.17) we conclude that the energy detector suffers a performance loss equal to $\sqrt{n}$, when compared to the incoherent detector. As shown in the next subsection, these conclusions are not even good approximations of performance comparisons between the optimal Neyman-Pearson Detection schemes.

2.5.2 Detection Performance of the Neyman-Pearson Detectors

We can compare the performance of the Neyman-Pearson Detectors via the use of numerical methods which evaluate the probability of detection p_d , as a function of the variables α, ρ, and n. We first note that the incoherent and coherent detectors' performance can be expressed as a function of α and the product ρn. Let $\gamma = \rho n$ and let us then define $\gamma_{coh}(p_d, p_{fa})$ and $\gamma_{incoh}(p_d, p_{fa})$ as the smallest γ, which achieves the probability of detection p_d while maintaining the p_{fa}, respectively for the coherent and incoherent Neyman-Pearson Detectors. The ratio $\gamma_{incoh}(p_d, p_{fa})/\gamma_{coh}(p_d, p_{fa})$ is similar to (Eq. 2.5.16) and represents the performance gain from knowing the initial phase offset, as compared to the initial phase offset being completely unknown. The energy detectors' performance depends on α, ρ, and n. Define $\rho_{energy}(p_d, p_{fa}, n)$ as the smallest ρ, which achieves probability of detection p_d while maintaining the p_{fa}, assuming that the number of samples in of the Neyman-Pearson Energy detector is n. Define $n_{energy}(p_d, p_{fa}, \rho)$ and $n_{incoh}(p_d, p_{fa}, \rho)$ similarly. The ratios $\rho_{energy}(p_d, p_{fa}, n)/\gamma_{incoh}(p_d, p_{fa})$ and $n_{energy}(p_d, p_{fa}, \rho)/n_{incoh}(p_d, p_{fa}, \rho)$ are analogous to (Eq. 2.5.17) and represent the performance gain between the incoherent and energy detection schemes.

Figure 2-2 shows $\gamma_{incoh}(p_d, p_{fa})/\gamma_{coh}(p_d, p_{fa})$. From this figure it is clear that incoherent detection is within 1.5 dB the performance of coherent detection for a $p_{fa} \leq 0.001$, as compared at any reasonable probability of detection. This is in contrast to the 3 dB performance difference derived using just first and second order statistics.

Figure 2-3, Figure 2-4, and Figure 2-5 show $\rho_{energy}(p_d, p_{fa}, n)/\gamma_{incoh}(p_d, p_{fa})$ for $n = 10, 100$, and 1000 respectively. From these figures it is clear that energy detection performs significantly better than the $\sqrt{n}$ performance loss (in terms of signal-to-noise ratio) derived from just first and second order statistics at all the p_{fa} and p_d values covered in these figures.

Finally, Figure 2-6, Figure 2-7, and Figure 2-8 show $n_{energy}(p_d, p_{fa}, \rho)/n_{incoh}(p_d, p_{fa}, \rho)$ for $p_{fa} = 1e-1, 1e-3$ and $1e-5$ respectively. From these figures it is clear that energy detection performs significantly better than the $\sqrt{n}$ performance loss (in terms of samples needed) derived from just first and second order statistics at all the p_d values covered in these figures.

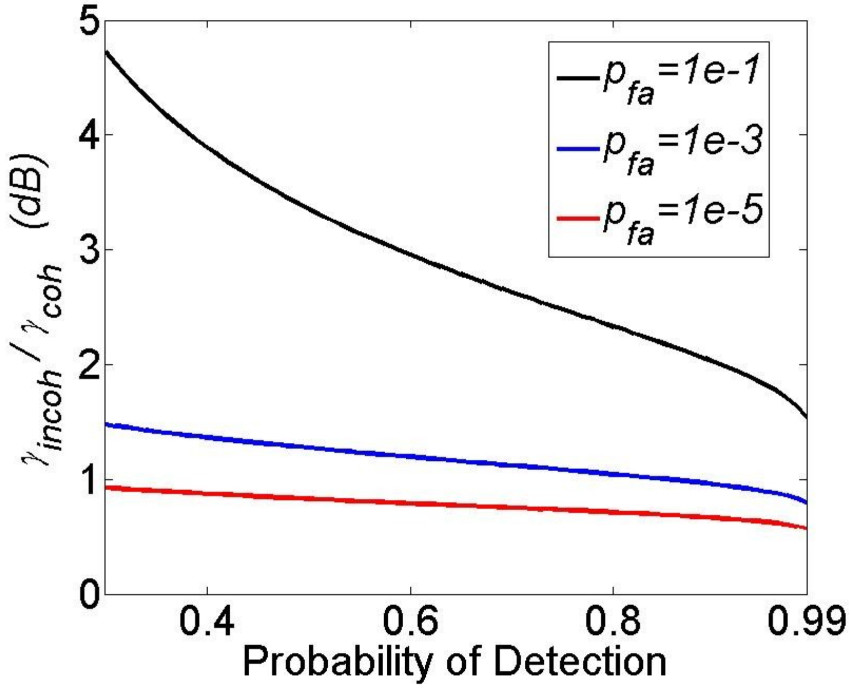

Figure 2-2: $\gamma_{incoh}/\gamma_{coh}$ as function of p_d for various p_{fa}.

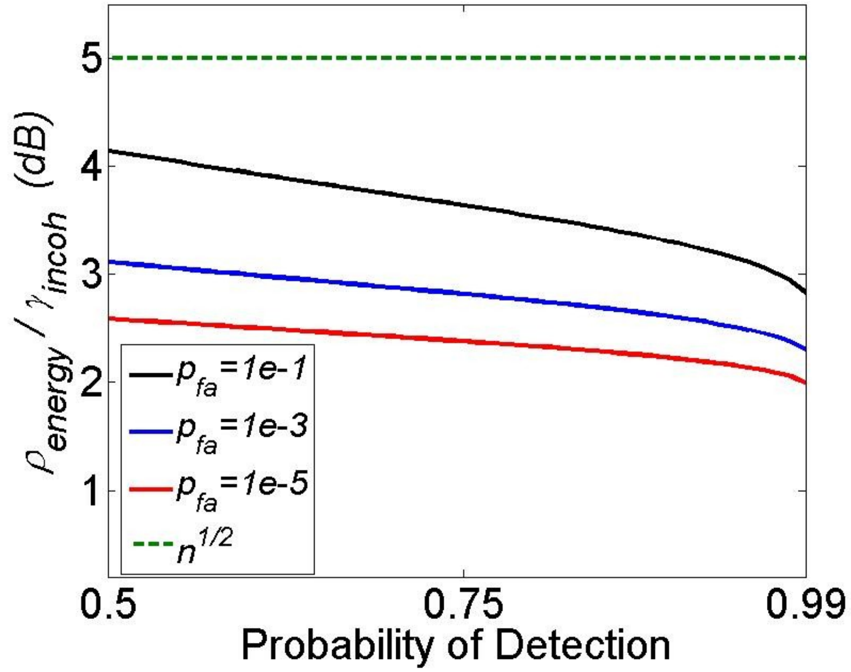

Figure 2-3: $\rho_{energy}/\gamma_{incoh}$ as function of p_d for various p_{fa} with $n = 10$.

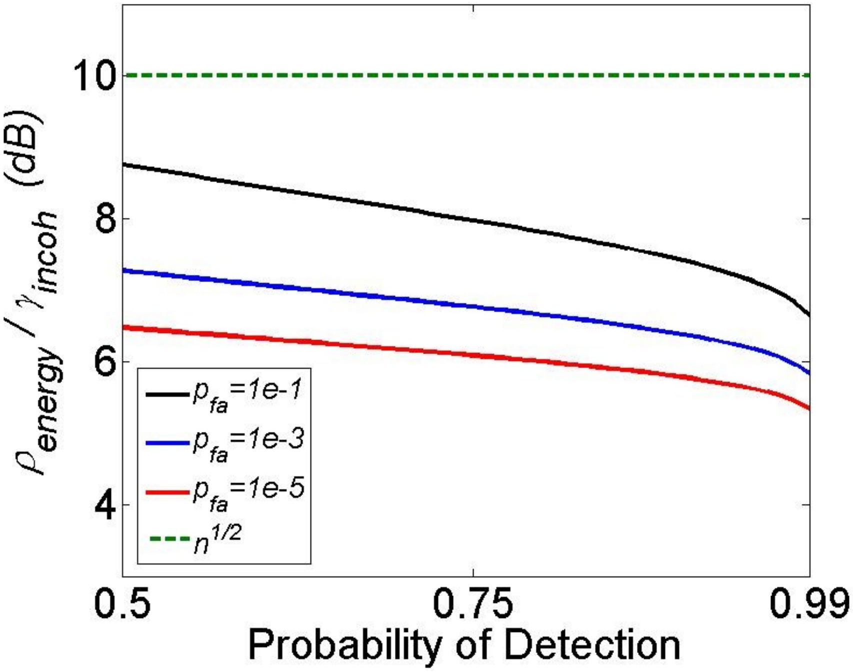

Figure 2-4: $\rho_{energy}/\gamma_{incoh}$ as function of p_d for various p_{fa} with $n = 100$.

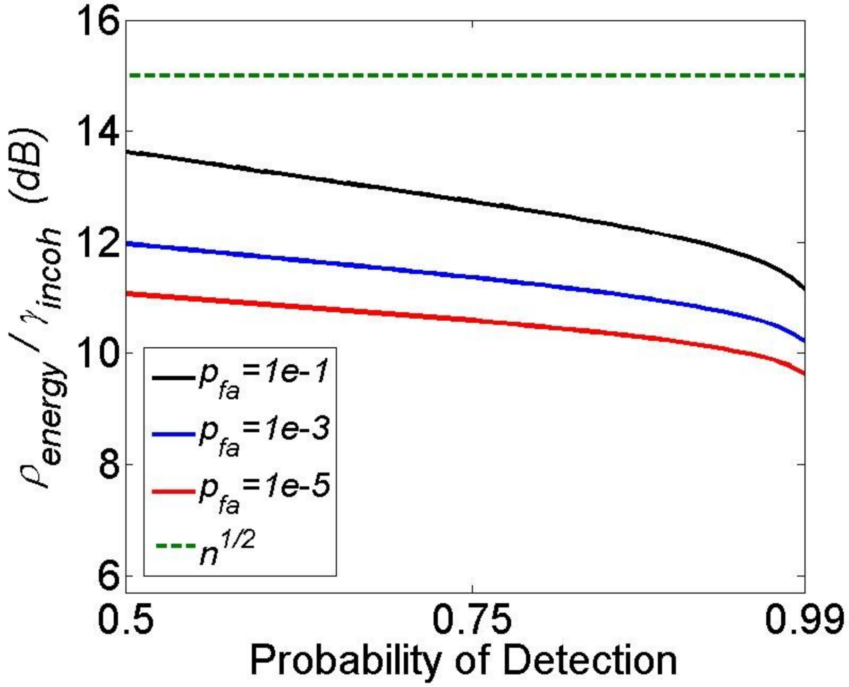

Figure 2-5: $\rho_{energy}/\gamma_{incoh}$ as function of p_d for various p_{fa} with $n = 1000$.

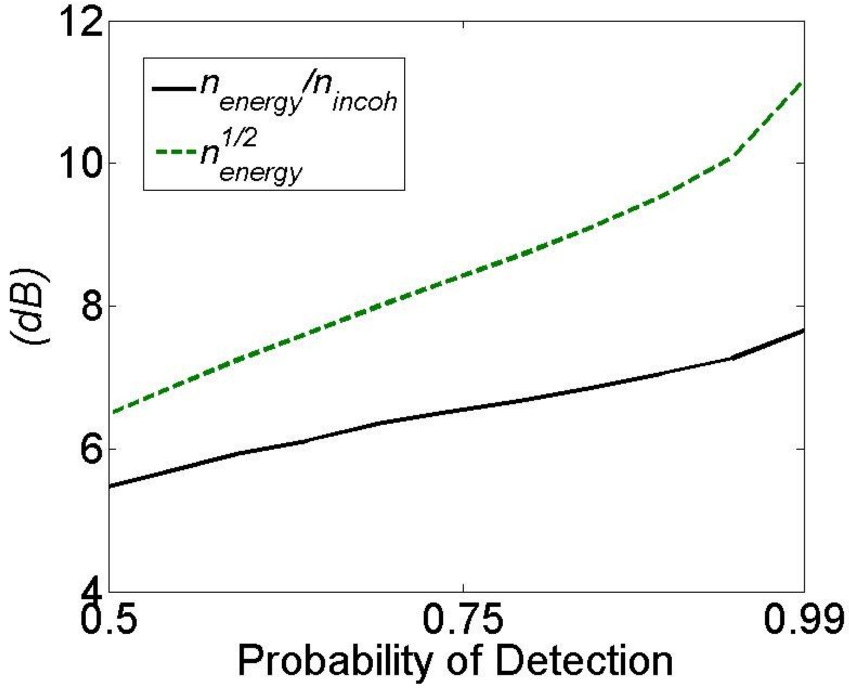

Figure 2-6: n_{energy}/n_{incoh} as a function of p_d with $p_{fa} = 1e - 1$ and $\rho = -5\ dB$.

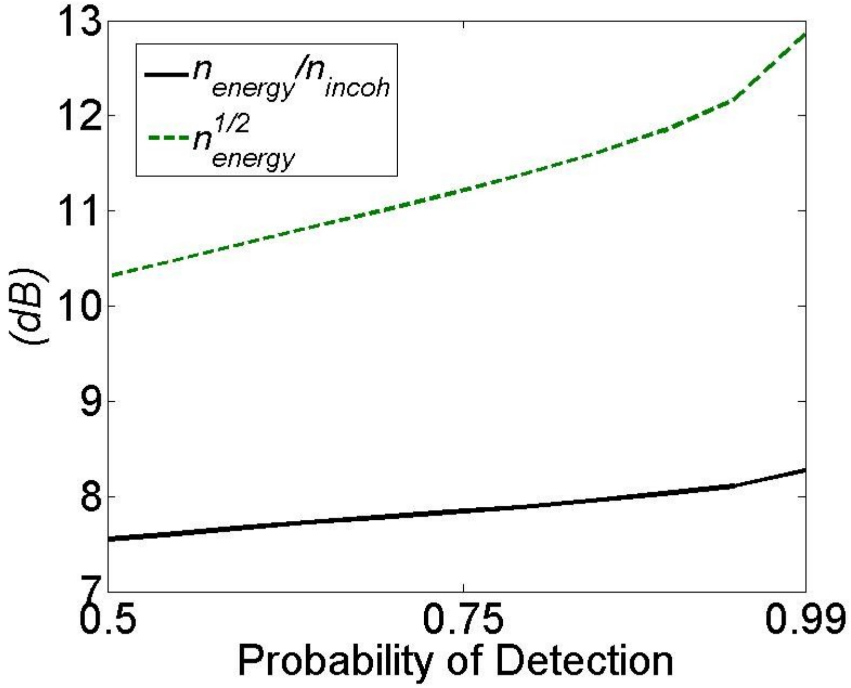

Figure 2-7: n_{energy}/n_{incoh} as a function of p_d with $p_{fa} = 1e-3$ and $\rho = -5\ dB$.

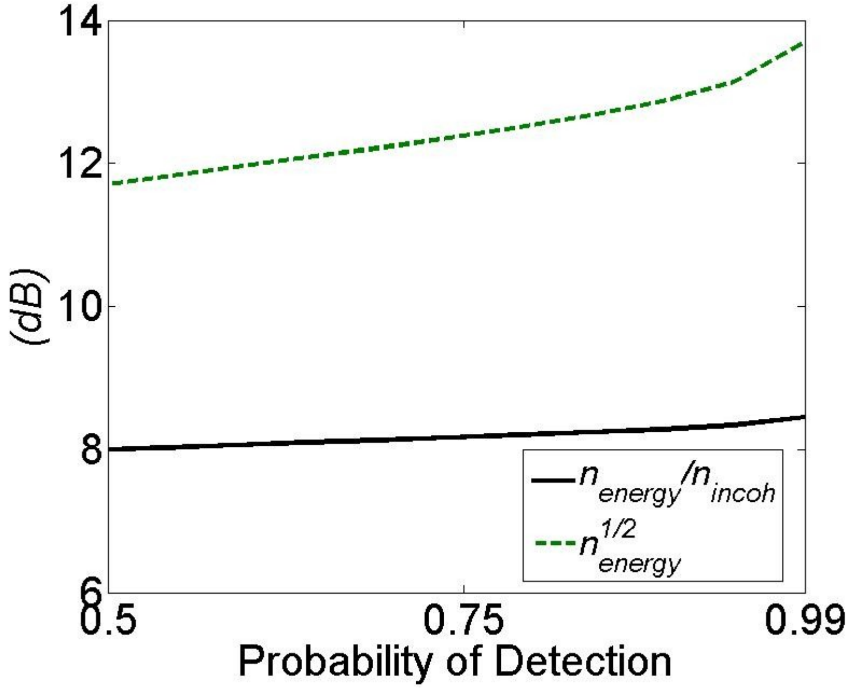

Figure 2-8: n_{energy}/n_{incoh} as a function of p_d with $p_{fa} = 1e-5$ and $\rho = -5\ dB$.

3 Robust Detection

3.1 Detection when the waveform and the initial phase are known

3.1.1 Case where contamination occurs over data pairs

Let $\Im_1$ and $\Im_o$ be two disjoint classes of n-dimensional density functions, and let $\Im_1$ and $\Im_o$ represent respectively the hypotheses H_1 and H_o. Assume that x^n is generated by a single density function in $\Im_1 \cup \Im_o$. Furthermore, consider that the two classes are Huber classes defined as

$$\Im_o = \{f(x_i, y_i) = (1 - \varepsilon_o) f_o(x_i, y_i) + \varepsilon_o h(x_i, y_i) \qquad \text{(Eq. 3.1.1)}$$

$$\Im_1 = \{f(x_i, y_i) = (1 - \varepsilon_1) f_{1,o}(x_i, y_i) + \varepsilon_1 h(x_i, y_i) \qquad \text{(Eq. 3.1.2)}$$

where

$$f_o(x_i, y_i) = \frac{1}{2\pi N_o} \exp\left(\frac{1}{2N_o}\left(x_i^2 + y_i^2\right)\right) \qquad \text{(Eq. 3.1.3)}$$

$$\begin{aligned} f_{1,i}(x_i, y_i) = \frac{1}{2\pi N_o} \exp\bigg(&\frac{1}{2N_o}\left(x_i - \sqrt{E} cos(p_i + \theta_o)\right)^2 \\ &+ \frac{1}{2N_o}\left(y_i - \sqrt{E} sin(p_i + \theta_o)\right)^2\bigg) \end{aligned} \qquad \text{(Eq. 3.1.4)}$$

and where $\varepsilon_o, \varepsilon_1$ are two given real constants whose values lie in $(0,1)$. The classes above model the cases where a sequence of data is generated by a corresponding nominal process, and contaminations may appear independently per data pair (x_i, y_i) with probabilities ε_o for class $\Im_o$ and ε_1 for class $\Im_1$. We later consider the case where the contamination may occur independently per datum x_i or y_i.

Based on the results derived by Huber, as presented in *Detection and Estimation* (169-74) [3], and extended in Appendix A, we have

$$\gamma_i^*(x_i, y_i) = \log\left(\frac{f_{1,i}^*(x_i, y_i)}{f_{0,i}^*(x_i, y_i)}\right) \tag{Eq. 3.1.5}$$

$$= \begin{cases} \ln c_{1,i} + \ln\left(\dfrac{1-\varepsilon_1}{1-\varepsilon_0}\right) & x_i, y_i \colon \beta_i(x_i, y_i) \leq \ln c_{1,i} \\ \ln\left(\dfrac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \beta_i(x_i, y_i) & x_i, y_i \colon \ln c_{1,i} < \beta_i(x_i, y_i) < \ln c_{o,i} \\ \text{lln}\, c_{0,i} + \ln\left(\dfrac{1-\varepsilon_1}{1-\varepsilon_0}\right) & x_i, y_i \colon \beta_i(x_i, y_i) \geq \ln c_{o,i} \end{cases} \tag{Eq. 3.1.6}$$

where

$$\begin{aligned} \beta_i(x_i, y_i) = {} & \frac{\sqrt{E}cos(p_i + \theta_o)}{2N_0}\left(2x_i - \sqrt{E}cos(p_i + \theta_o)\right) \\ & + \frac{\sqrt{E}sin(p_i + \theta_o)}{2N_0}\left(2y_i - \sqrt{E}sin(p_i + \theta_o)\right) \end{aligned} \tag{Eq. 3.1.7}$$

Let us define

$$k_i = x_i cos(p_i + \theta_o) + y_i sin(p_i + \theta_o) \tag{Eq. 3.1.8}$$

Then

$$\beta_i(x_i, y_i) = \beta_i(k_i) = \frac{\sqrt{E}}{N_o}k_i - \frac{E}{2N_o} = \frac{\sqrt{E}}{2N_o}\left(2k_i - \sqrt{E}\right) \tag{Eq. 3.1.9}$$

The statistics of k_i under $f_{1,i}(k_i), f_0(k_i)$ are the same for all i and therefore the statistics of k_i under $f_{1,i}^*(k_i), f_{0,i}^*(k_i)$ are the same for all i. Thus, from this point on, we can drop the index i from $\gamma_i^*, c_{1,i}, c_{0,i}, \beta_i$. Let us then define

$$z(k_i) = \frac{N_o}{\sqrt{E}}\gamma^*(k_i) - \frac{N_o}{\sqrt{E}}\ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \frac{\sqrt{E}}{2} \qquad \text{(Eq. 3.1.10)}$$

$$d_1 = \frac{N_o}{\sqrt{E}}\ln c_1 + \frac{\sqrt{E}}{2} \qquad \text{(Eq. 3.1.11)}$$

$$d_0 = \frac{N_o}{\sqrt{E}}\ln c_0 + \frac{\sqrt{E}}{2} \qquad \text{(Eq. 3.1.12)}$$

Then

$$z(k_i) = \begin{cases} \mathrm{d}_1 & k_i \le d_1 \\ k_i & d_1 < k_i < d_0 \\ \mathrm{d}_0 & k_i \ge d_0 \end{cases} \qquad \text{(Eq. 3.1.13)}$$

The analysis has been simplified down to the case where a deterministic scalar constant signal may be transmitted in additive white Gaussian noise and follows Example 6.3.1 presented in *Detection and Estimation*. The key to this simplification lies in the assumption that contamination occurs over data pairs (x_i, y_i). We include some of the key results and equations of the analysis found in *Detection and Estimation* as future sections will reference them. The reader should refer to *Detection and Estimation* (p. 175-85) for more details [3].

We will focus our analysis on the case where $\varepsilon_o = \varepsilon_1 = \varepsilon < 0.5$.

The robust decision rule is

$$\delta^*(k^n) = \begin{cases} 1 & k^n : n^{-1}\sum_{i=1}^{n} z(k_i) > \lambda^* \\ r^* & k^n : n^{-1}\sum_{i=1}^{n} z(k_i) = \lambda^* \\ 0 & k^n : n^{-1}\sum_{i=1}^{n} z(k_i) < \lambda^* \end{cases} \qquad \text{(Eq. 3.1.14)}$$

where $z(k_i)$ is as in (Eq. 3.1.13) and where λ^* and r^* are such that

$$\Pr\left\{n^{-1}\sum_{i=1}^{n} z(k_i) > \lambda^* \middle| f_0^*\right\} + r^* \Pr\left\{n^{-1}\sum_{i=1}^{n} z(k_i) = \lambda^* \middle| f_0^*\right\} = \alpha \qquad \text{(Eq. 3.1.15)}$$

The least favorable density functions f_0^* and f_1^* are

$$f_0^*(k) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{k^2}{2N_o}\right) & k: k < d_0 \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-\left(k-\sqrt{E}\right)^2}{2N_o} + \dfrac{E}{2N_o} - \dfrac{\sqrt{E}}{N_o} d_o\right) & k: k \geq d_0 \end{cases} \qquad \text{(Eq. 3.1.16)}$$

$$f_1^*(k) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{\left(k-\sqrt{E}\right)^2}{2N_o}\right) & k: k > d_1 \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-k^2}{2N_o} - \dfrac{E}{2N_o} + \dfrac{\sqrt{E}}{N_o} d_1\right) & k: k \leq d_1 \end{cases} \qquad \text{(Eq. 3.1.17)}$$

where

$$2^{-1}\sqrt{E} \geq d = d_1 = -d_o + \sqrt{E} \qquad \text{(Eq. 3.1.18)}$$

and d is the unique solution of

$$(1-\varepsilon)\left[\Phi\left(-\frac{d}{\sqrt{N_o}} + \sqrt{\frac{E}{N_o}}\right) + \exp\left(\frac{d\sqrt{E}}{N_o} - \frac{E}{2N_o}\right)\Phi\left(\frac{d}{\sqrt{N_o}}\right)\right] = 1 \qquad \text{(Eq. 3.1.19)}$$

where there is the solution to (Eq. 3.1.19), which results in a valid d if

$$\varepsilon \leq \varepsilon_b = 1 - 2^{-1}\left[\Phi\left(\frac{\sqrt{E}}{2\sqrt{N_o}}\right)\right]^{-1} \qquad \text{(Eq. 3.1.20)}$$

Since the variables $\{K_i\}$ are independent and identically distributed when conditioned on any f in $\mathfrak{I}_1 \cup \mathfrak{I}_o$, the random variables $\{\beta_i = z(K_i)\}$ are too. The random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$ is then asymptotically Gaussian. Let us denote

$$\rho = \sqrt{E}/\sqrt{N_o} \qquad \text{(Eq. 3.1.21)}$$

$$c(\varepsilon) = d/\sqrt{N_o} \qquad \text{(Eq. 3.1.22)}$$

Denote $m_{f_o^*}$ and $v_{f_o^*}^2$ as the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$ as n gets large, conditioned on f_o^*. We find

$$\begin{aligned} m_{f_o^*} = (1-\varepsilon)\sqrt{N_o}\{&\Phi[c(\varepsilon)][c(\varepsilon) + [\rho - c(\varepsilon)]exp[\rho c(\varepsilon) - 2^{-1}\rho^2]] \\ &+ \phi[c(\varepsilon)] - \phi[c(\varepsilon) - \rho]\} \end{aligned} \qquad \text{(Eq. 3.1.23)}$$

$$\begin{aligned} v_{f_o^*}^2 = (1-\varepsilon)N_o n^{-1}\Big\{&\Phi[c(\varepsilon)]\Big[c^2(\varepsilon) + \big(\rho - c(\varepsilon)\big)^2 \exp(pc(\varepsilon) - 2^{-1}\rho^2) \\ &- 1\Big] + \Phi[\rho - c(\varepsilon)] + c(\varepsilon)\phi[c(\varepsilon)] + [c(\varepsilon) - \rho]\phi[\rho - c(\varepsilon)] \\ &- (1-\varepsilon)^{-1}N_o^{-1}m_{f_o^*}^2\Big\} \end{aligned} \qquad \text{(Eq. 3.1.24)}$$

For asymptotically large values of n, the constant r^* in (Eq. 3.1.15) can be taken equal to zero and we obtain

$$\lambda^* = m_{f_0^*} + v_{f_0^*}\,\Phi^{-1}(1-\alpha) \qquad \text{(Eq. 3.1.25)}$$

Denote m_{f_1} and $v_{f_1}^2$ as the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$ as n gets large, conditioned on f_1. We find

$$m_{f_1} = \sqrt{N_o}\left[\rho - \phi\left(c(\varepsilon)\right) + \phi(c(\varepsilon) - \rho) + (c(\varepsilon) - \rho)\Phi(c(\varepsilon) - \rho) - c(\varepsilon)\Phi(c(\varepsilon))\right] \qquad \text{(Eq. 3.1.26)}$$

$$v_{f_1}^2 = n^{-1}N_o\{c^2(\varepsilon) + [1 + \rho^2 - c^2(\varepsilon)]\Phi[\rho - c(\varepsilon)] + (c^2(\varepsilon) - 2\rho c(\varepsilon) - 1)\Phi[c(\varepsilon)] - [-c(\varepsilon) + 2\rho]\phi[c(\varepsilon)] + [\rho + c(\varepsilon)]\phi[\rho - c(\varepsilon)] - N_o^{-1}m_{f_1}^2\} \qquad \text{(Eq. 3.1.27)}$$

Denote by m_{g_m} and $v_{g_m}^2$ respectively the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$, as n gets large, conditioned on g_m(K) , where g_m is Gaussian with variance N_o and mean m. Define

$$\mu = m/\sqrt{N_o} \qquad \text{(Eq. 3.1.28)}$$

We find

$$m_{g_m} = \sqrt{N_o}\{\mu + [\rho - \mu - c(\varepsilon)]\Phi(c(\varepsilon) - \rho + \mu) + [c(\varepsilon) - \mu]\Phi(c(\varepsilon) - \mu) + \phi(c(\varepsilon) - \mu) - \phi\left(\rho - \mu - c(\varepsilon)\right)\} \qquad \text{(Eq. 3.1.29)}$$

$$v_{g_m}^2 = n^{-1}N_o\{1 + \mu^2 + [c^2(\varepsilon) - \mu^2 - 1]\Phi[c(\varepsilon) - \mu] + [(c(\varepsilon) - \rho)^2 - \mu^2 - 1\,]\Phi[c(\varepsilon) - \rho + \mu] + [c(\varepsilon) + \mu]\phi[c(\varepsilon) - \mu] + [c(\varepsilon) - \rho - \mu]\phi[\rho - \mu - c(\varepsilon)] - N_o^{-1}m_{g_m}^2\} \qquad \text{(Eq. 3.1.30)}$$

For

$$f = (1-\varepsilon)f_1 + \varepsilon g_m \qquad \text{(Eq. 3.1.31)}$$

we obtain

$$m_f = (1-\varepsilon)m_{f_1} + \varepsilon m_{g_m} \qquad \text{(Eq. 3.1.32)}$$

$$v_f^2 = (1-\varepsilon)v_{f_1}^2 + \varepsilon v_{g_m}^2 + n^{-1}\varepsilon(1-\varepsilon)\left[m_{f_1} - m_{g_m}\right]^2 \qquad \text{(Eq. 3.1.33)}$$

Let us denote $P_n(f)$ the power induced by the robust decision rule in (Eq. 3.1.15) at the density function f in (Eq. 3.1.31). Due to (Eq. 3.1.25), we conclude: for $n \to \infty$,

$$P_n(f) = \Phi\left(\frac{m_f - m_{f_0^*} - v_{f_0^*}\Phi^{-1}(1-\alpha)}{v_f}\right) \qquad \text{(Eq. 3.1.34)}$$

3.1.2 Case where contamination occurs per datum

For this case, where the waveform is completely known, treat each complex pair as an individual sample. Define

$$u_1^{2n} = (x_1, y_1, x_2, x_3, \dots., x_n, y_n) \qquad \text{(Eq. 3.1.35)}$$

Also define

$$\begin{aligned} &h_1^{2n} \\ &= (\sqrt{\mathrm{E}}\cos\big(\theta_o + p(1)\big), \sqrt{\mathrm{E}}\sin\big(\theta_o \\ &\qquad + p(1), \dots, \sqrt{\mathrm{E}}\cos\big(\theta_o + p(n)\big), \sqrt{\mathrm{E}}\sin\big(\theta_o + p(n)\big)\big) \end{aligned} \qquad \text{(Eq. 3.1.36)}$$

In other words, h_1^{2n} is the known signal. With this new notation,

$$\frac{f_1(u_i)}{f_0(u_i)} = \exp\left(\frac{h_i}{2N_0}(2u_i - h_i)\right) \quad \text{(Eq. 3.1.37)}$$

Our problem can be stated as follows. Let $\mathfrak{F}_1$ and $\mathfrak{F}_o$ be two disjoint classes of n-dimensional density functions, and let $\mathfrak{F}_1$ and $\mathfrak{F}_o$ represent respectively hypotheses H_1 and H_o. Assume that u^{2n} is generated by a single density function in $\mathfrak{F}_1 \cup \mathfrak{F}_o$. Furthermore, consider that the two classes are Huber classes, defined as

$$\mathfrak{F}_o = \{f(u_i) = (1 - \varepsilon_o)f_o(u_i) + \varepsilon_o h(u_i) \quad \text{(Eq. 3.1.38)}$$

$$\mathfrak{F}_1 = \{f(u_i) = (1 - \varepsilon_1)f_{1,i}(u_i) + \varepsilon_1 h(u_i) \quad \text{(Eq. 3.1.39)}$$

where

$$f_o(u_i) = \frac{1}{\sqrt{2\pi N_o}}\exp\left(\frac{1}{2N_o}\left(u_i^2\right)\right) \quad \text{(Eq. 3.1.40)}$$

$$f_{1,i}(u_i) = \frac{1}{\sqrt{2\pi N_o}}\exp\left(\frac{1}{2N_o}(u_i - h_i)^2\right) \quad \text{(Eq. 3.1.41)}$$

and where $\varepsilon_o, \varepsilon_1$ are two given real constants whose values lie in $(0,1)$. The classes above model the cases where a sequence of data is generated by a corresponding nominal process, and contaminations may appear independently per datum (u_i) with probabilities ε_o for class $\mathfrak{F}_o$ and ε_1 for class $\mathfrak{F}_1$.

Based on the results derived by Huber, as presented in *Detection and Estimation* (169-74) [3], and extended in Appendix A, we have

$$\log\left(\frac{f_{1,i}(u_i)}{f_{0,i}(u_i)}\right) = \frac{h_i}{2N_0}(2u_i - h_i) \quad \text{(Eq. 3.1.42)}$$

$$\gamma_i^*(u_i) = \log\left(\frac{f_{1,i}^*(u_i)}{f_{0,i}^*(u_i)}\right) = \qquad \text{(Eq. 3.1.43)}$$

$$\begin{cases} \ln c_{1,i} + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & u_i\text{: } \beta_i(u_i) \le \ln c_{1,i} \\ \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \beta_i(x_i, y_i) & u_i\text{: } \ln c_{1,i} < \beta_i(u_i) < \ln c_{o,i} \\ \ln c_{0,i} + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & u_i\text{: } \beta_i(u_i) \ge \ln c_{o,i} \end{cases} \qquad \text{(Eq. 3.1.44)}$$

where

$$\beta_i(u_i) = \frac{h_i}{2N_0}(2u_i - h_i) \qquad \text{(Eq. 3.1.45)}$$

Let us define

$$z_i(u_i) = \frac{N_o}{h_i}\gamma_i^*(u_i) - \frac{N_o}{h_i}\ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \frac{h_i}{2} \qquad \text{(Eq. 3.1.46)}$$

$$d_{1,i} = \frac{N_o}{h_i}\ln c_{1,i} + \frac{h_i}{2} \qquad \text{(Eq. 3.1.47)}$$

$$d_{0,i} = \frac{N_o}{h_i}\ln c_{0,i} + \frac{h_i}{2} \qquad \text{(Eq. 3.1.48)}$$

Then, if $h_i > 0$

$$z_i(u_i) = \begin{cases} \mathrm{d}_{1,\mathrm{i}} & u_i \le d_{1,i} \\ u_i & d_1 < u_i < d_{0,i} \\ \mathrm{d}_{0,\mathrm{i}} & u_i \ge d_{0,i} \end{cases} \qquad \text{(Eq. 3.1.49)}$$

or if $h_i < 0$

$$z_i(u_i) = \begin{cases} \mathrm{d_{1,i}} & u_i \geq d_{1,i} \\ u_i & d_1 > \ u_i > d_{0,i} \\ \mathrm{d_{0,i}} & u_i \leq d_{0,i} \end{cases} \qquad \text{(Eq. 3.1.50)}$$

We will focus our analysis on the case where $\varepsilon_o = \varepsilon_1 = \varepsilon < 0.5$.

The robust decision rule takes the form

$$\delta^*(u^n) = \begin{cases} 1 \quad u^{2n}\colon 2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z_i(u_i) > \lambda^* \\ r^* \quad u^{2n}\colon 2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z_i(u_i) = \lambda^* \\ 0 \quad u^{2n}\colon 2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z_i(u_i) < \lambda^* \end{cases} \qquad \text{(Eq. 3.1.51)}$$

where $z_i(u_i)$ are as in (Eq. 3.1.50) and where λ^* and r^* are such that

$$\Pr\left\{2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z(u_i) > \lambda^* \middle| f_0^*\right\} + r^* \Pr\left\{2^{-1}n^{-1}\sum_{i=1}^{n} h_i z(u_i) = \lambda^* \middle| f_0^*\right\} = \alpha \qquad \text{(Eq. 3.1.52)}$$

The least favorable density functions $f_{0,i}^*$ and $f_{1,i}^*$, if $h_i > 0$ are

$$f_{0,i}^*(u_i) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{u_i^2}{2N_o}\right) & u_i\colon u_i < d_{0,i} \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-(u_i-h_i)^2}{2N_o} + \dfrac{h_i^2}{2N_o} - \dfrac{h_i}{N_o} d_{o,i}\right) & u_i\colon u_i \geq d_{0,i} \end{cases} \qquad \text{(Eq. 3.1.53)}$$

$$f_{1,i}^*(u_i) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{(u_i-h_i)^2}{2N_o}\right) & u_i\colon u_i > d_{1,i} \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-u_i^2}{2N_o} - \dfrac{h_i^2}{2N_o} + \dfrac{h_i}{N_o} d_{1,i}\right) & u_i\colon u_i \leq d_{1,i} \end{cases} \qquad \text{(Eq. 3.1.54)}$$

and if $h_i > 0$ are

$$f_{0,i}^{*}(u_i) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{u_i^2}{2N_o}\right) & u_i\colon u_i > d_{0,i} \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-(u_i-h_i)^2}{2N_o} + \dfrac{h_i^2}{2N_o} - \dfrac{h_i}{N_o} d_{o,i}\right) & u_i\colon u_i \le d_{0,i} \end{cases} \qquad \text{(Eq. 3.1.55)}$$

$$f_{1,i}^{*}(u_i) = \begin{cases} \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(-\dfrac{(u_i-h_i)^2}{2N_o}\right) & u_i\colon u_i < d_{1,i} \\ \dfrac{(1-\varepsilon)}{\sqrt{2\pi N_o}} \exp\left(\dfrac{-u_i^2}{2N_o} - \dfrac{h_i^2}{2N_o} + \dfrac{h_i}{N_o} d_{1,i}\right) & u_i\colon u_i \ge d_{1,i} \end{cases} \qquad \text{(Eq. 3.1.56)}$$

where

$$d_i = d_1 = -d_o + h_i \qquad \text{(Eq. 3.1.57)}$$

and $\{d_i\}$ are the solutions of

$$(1-\varepsilon)\left[\Phi\left(-\frac{d_i}{\sqrt{N_o}} + \frac{h_i}{\sqrt{N_o}}\right) + \exp\left(\frac{d_i h_i}{N_o} - \frac{h_i^2}{2N_o}\right)\Phi\left(\frac{d_i}{\sqrt{N_o}}\right)\right] = 1;\ \ h_i\colon h_i > 0 \qquad \text{(Eq. 3.1.58)}$$

$$(1-\varepsilon)\left[\Phi\left(\frac{d_i}{\sqrt{N_o}} - \frac{h_i}{\sqrt{N_o}}\right) + \exp\left(\frac{d_i h_i}{N_o} - \frac{h_i^2}{2N_o}\right)\Phi\left(\frac{-d_i}{\sqrt{N_o}}\right)\right] = 1;\ \ h_i\colon h_i < 0 \qquad \text{(Eq. 3.1.59)}$$

From (Eq. 3.1.51), we observe that the robust receiver performs a truncation per datum in a similar fashion as that of the robust receiver found in section 3.1.1, except that the function which performs truncation changes here for each sample, according to h_i. Also, (Eq. 3.1.49) and (Eq. 3.1.50) require that $|d_1| \le |d_o|$. Due to this requirement, we could have the case where some data samples have small enough $|h|$ so that no valid solution to (Eq. 3.1.58) or (Eq. 3.1.59) exists. Those data samples where no solution exists represent a case where $\mathfrak{I}_{1,i}$ and $\mathfrak{I}_{o,i}$ overlap. For our overall detection, such samples do not help our decision rule, and in the analysis that follows we will assign $z_i(u_i) = 0$ to all samples where no valid solution exist.

The variables $\{U_i\}$ are independent when conditioned on any f in $\mathfrak{I}_1 \cup \mathfrak{I}_o$. Let us denote a subset of $\{U_i\}$ as $\{U_{i,L} : U_i \; where \; L \le h_i < L + dh\}$ where dh is arbitrarily small. The members in subset $\{U_{i,L}\}$ are independent and identically distributed when conditioned on any f in $\mathfrak{I}_1 \cup \mathfrak{I}_o$, and so are the random variables $\{Y_{i,L} = z_L(U_{i,L})\}$. The random variable $n_L^{-1} \sum_{i=1}^{n_L} h_i z_L(U_{i,L})$, where n_L is the number of elements inside the subset, is then asymptotically Gaussian. Finally, $2^{-1}n^{-1} \sum_{i=1}^{2n} h_i z \; (U_{i,L})$ is a linear combination of asymptotically Gaussian random variables; it is therefore asymptotically Gaussian.

$$m_{f_0^*} = E\{2^{-1}n^{-1} \sum_{i=1}^{n} h_i z(U) | f_o^* \} = E\{\int_{-\sqrt{E}}^{\sqrt{E}} h_i z(U,h)\omega(h) dh \, | f_o^* \} \qquad \text{(Eq. 3.1.60)}$$

where $\omega(h)dh$ is the relative frequency of $\{U_i\}$ such that $h \le h_i < h + dh$. For a specified h_1^{2n} the integral in (Eq. 3.1.60) will reduce to a sum. Let us assume that, as n gets asymptotically large, $(\theta^*)d\theta^* = \frac{1}{2\pi}, -\pi < \theta^* \le \pi$; where $\theta^* = \theta_o + p(i)$ for x_i samples and $\theta^* = \frac{\pi}{2} - \theta_o - p(i)$ for y_i samples. Our model will be asymptotically correct, if the signal is such that, as n gets large, the relative occurrence of a sample falling into any phase band of width $d\theta^*$ is equally probable. Within the latter model, (Eq. 3.1.60) can be written as,

$$\begin{aligned} m_{f_0^*} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} E\{z(U,\theta^*) f_{o,\theta^*}^* \} \, d\theta^* \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} h_i E\{z(U, h = \sqrt{E}\cos(\theta^*) \, | f_{o,h}^*)\} \, d\theta^* \end{aligned} \qquad \text{(Eq. 3.1.61)}$$

Let us denote

$$\rho(h) = |h| / \sqrt{N_o} \qquad \text{(Eq. 3.1.62)}$$

$$c(\varepsilon, h) = d(h)/\sqrt{N_o} \qquad \text{(Eq. 3.1.63)}$$

With this notation, and based on (Eq. 3.1.23), we find

$$\begin{aligned}
&E\{z(U,h)|f^*_{o,h}\} \\
&= (1-\varepsilon)\sqrt{N_o}\{\mathit{\Phi}[c(\varepsilon,h)][c(\varepsilon,h) + [\rho(h) - c(\varepsilon,h)]exp[\rho(h)c(\varepsilon,h) \\
&\qquad\qquad - 2^{-1}\rho^2(h)]] \ + \phi[c(\varepsilon,h)] \qquad \text{(Eq. 3.1.64)} \\
&\qquad\qquad - \phi[c(\varepsilon,h) - \rho(h)]\}; \quad \varepsilon_b(h) > \varepsilon
\end{aligned}$$

$$E\{z(U,h)|f^*_{o,h}\} = 0; \quad \varepsilon_b(h) < \varepsilon$$

Define $h_b(\varepsilon) = \min\{|h|: \varepsilon_b(h) > \varepsilon\}$. Define $\theta_b(\varepsilon) = \cos^{-1}\left(\frac{h_b(\varepsilon)}{\sqrt{E}}\right)$. Due to symmetry, (Eq. 3.1.61) can be then simplified to the following form,

$$m_{f^*_0} = \frac{2}{\pi}\int_{\theta_b(\varepsilon)}^{\pi/2} \sqrt{E}\cos(\theta^*)\, E\{z(U, \sqrt{E}\cos(\theta^*)\,|f^*_{o,h})\}\, d\theta^* \qquad \text{(Eq. 3.1.65)}$$

Since we must find $c(\varepsilon, h)$ numerically, we must also evaluate (Eq. 3.1.65) numerically. In a similar fashion, and based on (Eq. 3.1.24), we find

$$v^2_{f^*_0} = (2n)^{-1}\frac{2}{\pi}\int_{\theta_b(\varepsilon)}^{\pi/2} E^2\cos^2(\theta^*)\, Var\{z(U, \sqrt{E}\cos(\theta^*)\,|f^*_{o,h})\}\, d\theta^* \qquad \text{(Eq. 3.1.66)}$$

$$\begin{aligned}
&Var\{z(U,h|f^*_{o,h})\} = \\
&(1-\varepsilon)N_o\left\{\mathit{\Phi}[c(\varepsilon,h)]\left[c^2(\varepsilon,h) + (\rho - c(\varepsilon,h))^2 \exp(p(h)c(\varepsilon,h) - \right.\right. \\
&2{-}1\rho2(h){-}1{+}\mathit{\Phi}\rho(h){-}c\varepsilon{,}h{+}c\varepsilon{,}h\phi c\varepsilon{,}h{+}c\varepsilon{,}h{-}\rho(h)\phi\rho(h){-}c\varepsilon{,}h{-}1{-}\varepsilon{-} \qquad \text{(Eq. 3.1.67)} \\
&1No{-}1E2zU{,}h|fo{,}h{*};
\end{aligned}$$

For asymptotically large values of n, the constant r^* in (Eq. 3.1.52) can be set at zero, resulting in

$$\lambda^* = m_{f_0^*} + v_{f_0^*}\,\Phi^{-1}(1-\alpha) \qquad \text{(Eq. 3.1.68)}$$

Asymptotically, $2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z\ \ (U_i)$ is Gaussian, when conditioned on $f = (1-\varepsilon)f_{1,i} + \varepsilon g_m$; where $g_{m,i}(u_i)$ is Gaussian with mean $mcos(\theta^*)$ and variance N_o. Define $\mu(h) = \ mh/(\sqrt{N_o E}) = mcos(\theta^*)/\sqrt{N_0}$. In (Eq. 3.2.26), (Eq. 3.2.27), (Eq. 3.2.29) and (Eq. 3.2.30) we can replace ρ with $\rho(h)$, $c(\varepsilon)$ with $c(\varepsilon, h)$, μ with $\mu(h)$ and n with $2n$ and integrate in a similar fashion, as in (Eq. 3.1.65) and (Eq. 3.1.66), to obtain the mean and variance (as $n \rightarrow \infty$) of $2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z\ \ (U_i)$; when conditioned on $f = (1-\varepsilon)f_{1,i} + \varepsilon g_{m,i}$.

$$m_{f_1} = \frac{2}{\pi}\int_0^{\theta_b(\varepsilon)} \sqrt{E}\cos(\theta^*)\,E\{z(U, \sqrt{E}\cos(\theta^*)\,|f_{1,h})\}\,d\theta^* \qquad \text{(Eq. 3.1.69)}$$

$$v_{f_1}^2 = (2n)^{-1}\frac{2}{\pi}\int_0^{\theta_b(\varepsilon)} E^2\cos^2(\theta^*)\,Var\{z(U, \sqrt{E}\cos(\theta^*)\,|f_{1,h})\}\,d\theta^* \qquad \text{(Eq. 3.1.70)}$$

$$m_{g_m} = \frac{2}{\pi}\int_0^{\theta_b(\varepsilon)} \sqrt{E}\cos(\theta^*)\,E\{z(U, \sqrt{E}\cos(\theta^*)\,|g_m)\}\,d\theta^* \qquad \text{(Eq. 3.1.71)}$$

$$v_{g_m}^2 = (2n)^{-1}\frac{2}{\pi}\int_0^{\theta_b(\varepsilon)} E^2\cos^2(\theta^*)\,Var\{z(U, \sqrt{E}\cos(\theta^*)\,|g_m)\}\,d\theta^* \qquad \text{(Eq. 3.1.72)}$$

$$\begin{aligned} E\{z(U, h|f_{1,h})\} = \sqrt{N_o}\,\big[&\rho(h) - \phi\big(c(\varepsilon, h)\big) + \phi(c(\varepsilon, h) - \rho(h)) \\ &+ (c(\varepsilon, h) - \rho(h))\,\Phi(c(\varepsilon, h) - \rho(h)) - c(\varepsilon, h)\,\Phi(c(\varepsilon, h))\big] \end{aligned} \qquad \text{(Eq. 3.1.73)}$$

$$
\begin{aligned}
&Var\left\{z\left(U,h|f_{1,h}\right)\right\} \\
&\quad = N_o\left\{c^2(\varepsilon,h)+[1+\rho^2(h)-c^2(\varepsilon,h)]\,\Phi[\rho(h)-c(\varepsilon,h)]\right. \\
&\quad + (c^2(\varepsilon,h)-2\rho(h)c(\varepsilon,h)-1)\,\Phi[c(\varepsilon,h)] \\
&\quad - [-(c(\varepsilon,h)+2\rho(h)]\phi[c(\varepsilon,h)] \\
&\quad + [\rho(h)+c(\varepsilon,h)]\phi[\rho(h)-c(\varepsilon,h)] \\
&\quad \left. - N_o^{-1}E^2\left\{z(U,h)|f_{1,h}\right\}\right\}
\end{aligned}
\tag{Eq. 3.1.74}
$$

$$
\begin{aligned}
E\{z(U,h|g_m)\} = \sqrt{N_o}\,\{&\mu(h) \\
&+ [\rho(h)-\mu(h)\ -c(\varepsilon,h)]\,\Phi(c(\varepsilon,h)-\rho(h)+\mu(h)\ ) \\
&+ [c(\varepsilon,h)-\mu(h)\ ]\,\Phi(c(\varepsilon,h)-\mu(h)\ ) \\
&+ \phi(c(\varepsilon,h)-\mu(h)\ )-\phi\big(\rho(h)-\mu(h)\ -c(\varepsilon,h)\big)\}
\end{aligned}
\tag{Eq. 3.1.75}
$$

$$
\begin{aligned}
&Var\left\{z\left(U,\sqrt{E}\cos(\theta^*)\,|g_m\right)\right\} \\
&\quad = N_o\{1+\mu^2(h) \\
&\quad + [c^2(\varepsilon,h)-\mu^2(h)-1]\,\Phi[c(\varepsilon,h)-\mu(h)\ ] \\
&\quad + [(c(\varepsilon,h)-\rho(h))^2-\mu^2(h)-1\ ]\,\Phi[c(\varepsilon,h)-\rho(h) \\
&\quad + \mu(h)\ ] + [c(\varepsilon,h)+\mu(h)\ ]\phi[c(\varepsilon,h)-\mu(h)\ ] \\
&\quad + [c(\varepsilon,h)-\rho(h)-\mu(h)\ ]\phi[\rho(h)-\mu(h)\ -c(\varepsilon,h)] \\
&\quad - N_o^{-1}E^2\{z(U,h)|g_m\}\}
\end{aligned}
\tag{Eq. 3.1.76}
$$

For

$$
f\ = (1-\varepsilon)f_{1,i}+\varepsilon g_m \tag{Eq. 3.1.77}
$$

we have

$$m_f = (1-\varepsilon)m_{f_1} + \varepsilon m_{g_m} \qquad \text{(Eq. 3.1.78)}$$

$$v_f^2 = (1-\varepsilon)v_{f_1}^2 + \varepsilon v_{g_m}^2 + n^{-1}\varepsilon(1-\varepsilon)\left[m_{f_1} - m_{g_m}\right]^2 \qquad \text{(Eq. 3.1.79)}$$

Let us denote by $P_n(f)$, the power induced by the robust decision rule in (Eq. 3.1.52), at the density function f in (Eq. 3.1.77). Due to (Eq. 3.1.68), we conclude: for $n \to \infty$,

$$P_n(f) = \Phi\left(\frac{m_f - m_{f_0^*} - v_{f_0^*}\Phi^{-1}(1-\alpha)}{v_f}\right) \qquad \text{(Eq. 3.1.80)}$$

3.1.3 Comparison of the robust cases with the optimal at the nominal model test

In section 2.2, we examined the Neyman-Pearson detection rule at the nominal model. The latter rule does not provide protection against data contamination represented by extreme outliers. In section 3.1.1, we examined the robust detection scheme in the case where contamination occurs per data pair(x_i, y_i). In section 3.1.2, we examined the robust detection scheme in the case where contamination occurs independently per datum. In this section, we will compare the asymptotic performance of all three detection schemes under all three contamination models (no contamination, contamination occurs per data pair, and contamination occurs per datum). We will use the notation $P_n^{X|Y}$ to denote the power of the detection scheme X under the contamination model Y, where X and Y can be $A, B, \text{or } C$ to respectively represent the cases of no contamination, contamination occurring per data pair, or contamination occurring per single datum. For example, $P_n^{B|C}$ represents the power of the robust detection rule derived assuming contamination occurring per data pair (B) under the model where contamination actually occurs per single datum (C). More specifically, for $X = B$, we will consider the power of the detection rule under f, as defined in (Eq. 3.1.31); and for $= C$, we will consider the power of the detection rule under f, as defined in (Eq. 3.1.77).

Comparison under the model of no contamination

Denote $\rho = \sqrt{\frac{E}{N_0}}$. From (Eq. 2.2.8), we have

$$P_n^{A|A} = \Phi\big(\rho\sqrt{n} + \Phi^{-1}(\alpha)\big) \qquad \text{(Eq. 3.1.81)}$$

Based on (Eq. 3.1.34), we have

$$P_n^{B|A} = \Phi\left(\frac{m_{f_1^B} - m_{f_0^{*B}} - v_{f_0^{*B}}\Phi^{-1}(1-\alpha)}{v_{f_1^B}}\right) \qquad \text{(Eq. 3.1.82)}$$

where $m_{f_0^{*B}}$, $v_{f_0^{*B}}$, $m_{f_1^B}$, and $v_{f_1^B}$ are defined in (Eq. 3.1.23), (Eq. 3.1.24), (Eq. 3.1.26), and (Eq. 3.1.27).

Based on (Eq. 3.1.80), we have

$$P_n^{C|A} = \Phi\left(\frac{m_{f_1^C} - m_{f_0^{*C}} - v_{f_0^{*C}}\Phi^{-1}(1-\alpha)}{v_{f_1^C}}\right) \qquad \text{(Eq. 3.1.83)}$$

where $m_{f_0^{*C}}$, $v_{f_0^{*C}}$, $m_{f_1^C}$, and $v_{f_1^C}$ are defined in (Eq. 3.1.65), (Eq. 3.1.66), (Eq. 3.1.69), and (Eq. 3.1.70). Figure 3-1 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon = 0.1, \alpha = 0.05$. Figure 3-2 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon = 0.01, \alpha = 0.05$. It can easily in both Figure 3-1 and Figure 3-2 that $P_n^{A|A} > P_n^{B|A} > P_n^{C|A}$ for all n values plotted. In other words, Figure 3-1 and Figure 3-2 show that the optimal detection scheme out performs the robust detection scheme if there are no outliers.

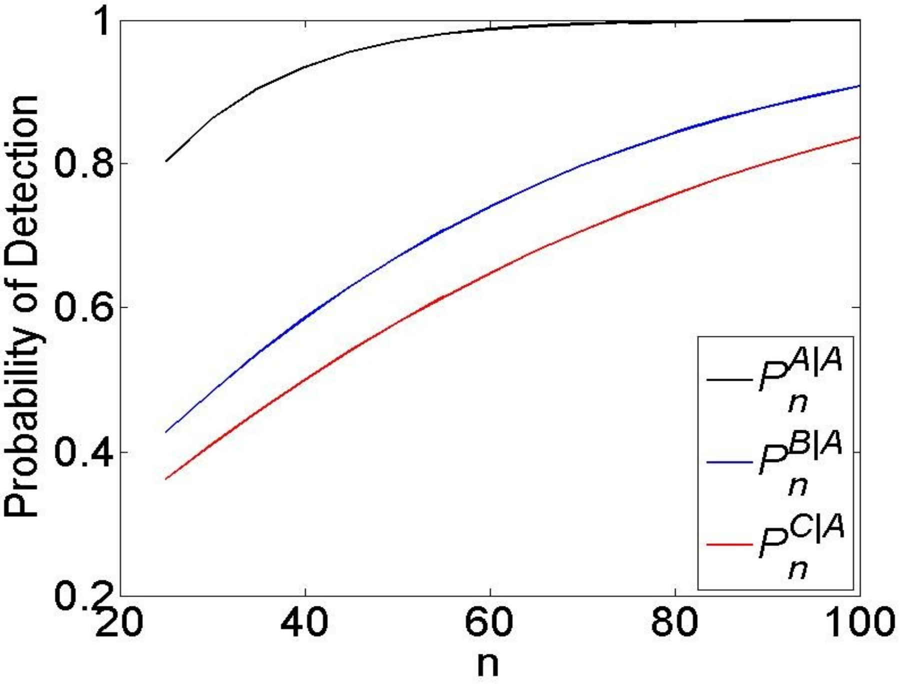

Figure 3-1: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.1, \alpha = 0.05$.

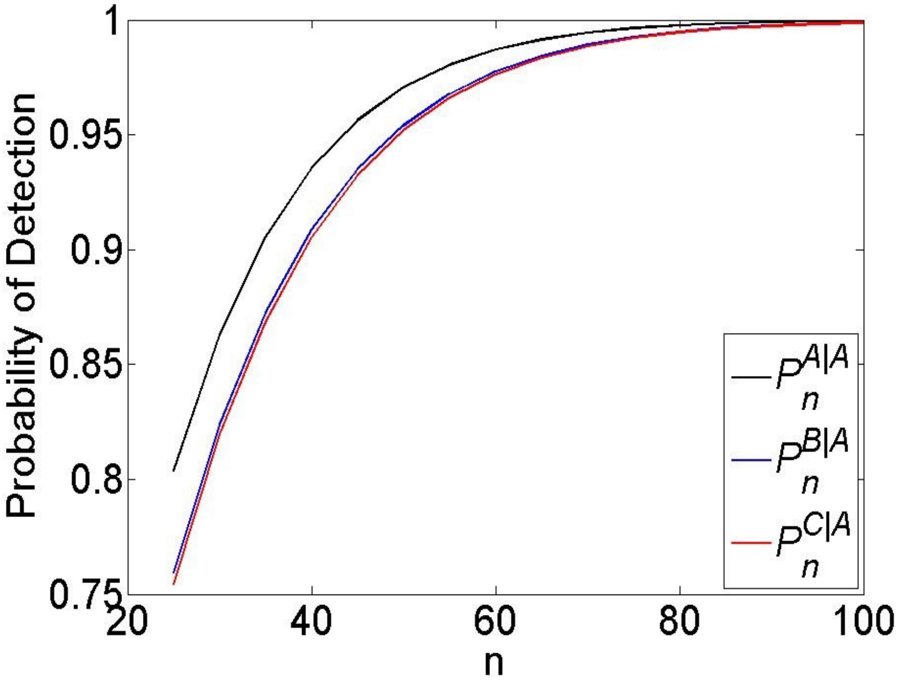

Figure 3-2: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.01, \alpha = 0.05$.

Comparison under the model of contamination per data pair

From Detection and Estimation (p. 184) [3], we obtain

$$P_n^{A|B} = \Phi\left(\frac{\sqrt{n}[(1-\varepsilon)\rho + \varepsilon\mu] + \Phi^{-1}(\alpha)}{\sqrt{1+\varepsilon(1-\varepsilon)(\rho-\mu)^2}}\right) \qquad \text{(Eq. 3.1.84)}$$

From (Eq. 3.1.34) we have

$$P_n^{B|B} = \Phi\left(\frac{m_{f^B} - m_{f_0^{*B}} - v_{f_0^{*B}}\Phi^{-1}(1-\alpha)}{v_{f^B}}\right) \qquad \text{(Eq. 3.1.85)}$$

where $m_{f_0^{*B}}$, $v_{f_0^{*B}}$, m_{f^B}, and v_{f^B} are defined in (Eq. 3.1.23),(Eq. 3.1.24),(Eq. 3.1.32), and (Eq. 3.1.33). To derive $P_n^{C|B}$, we consider density functions defined by

$$f(x_i) = (1-\varepsilon)f_{1,i}(x_i) + \varepsilon g_{m,x,i}(x_i) \qquad \text{(Eq. 3.1.86)}$$

$$f(y_i) = (1-z(i))f_{1,i}(y_i) + z(i)g_{m,y,i}(y_i) \qquad \text{(Eq. 3.1.87)}$$

where

$$z(i) = \begin{cases} 1;\ if\ x_i\ was\ a\ realization\ of\ g_{m,x,i}(x_i) \\ 0;\ if\ x_i\ was\ a\ realization\ of\ f_{1,i}(x_i) \end{cases} \qquad \text{(Eq. 3.1.88)}$$

These result in

$$f(x_i, y_i) = (1-\varepsilon)f_{1,i}(x_i, y_i) + \varepsilon g_{m,i}(x_i, y_i) \qquad \text{(Eq. 3.1.89)}$$

Furthermore, define $g_{m,x,i}(x_i)$ to be Gaussian with mean $mcos(\theta_o + p(i))$, and variance N_o, and $g_{m,y,i}(x_i)$ to be Gaussian with mean $msin(\theta_o + p(i))$ and variance N_o. We obtain

$$f(k_i) = (1 - \varepsilon)f_1(k_i) + \varepsilon g_m(k_i) \qquad \text{(Eq. 3.1.90)}$$

where $f_1(k_i)$ and $g_m(k_i)$ are the same densities as in Section 3.1.1. In other words, the densities defined by (Eq. 3.1.86) and (Eq. 3.1.87) result in the same density model assumed under contamination model B. We therefore can find $P_n^{C|B}$ as $n \to \infty$ by examining the statistics of $2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z^C(U_i)$ conditioned on (Eq. 3.1.86), and (Eq. 3.1.87) , where $z^C(u_i)$ is defined in (Eq. 3.1.49) and (Eq. 3.1.50). The density defined by (Eq. 3.1.86) and (Eq. 3.1.87) is similar to the density model assumed under contamination model C, except that the outliers always occur on consecutive values of u_i. The statistics of $2^{-1}n^{-1}\sum_{i=1}^{2n} h_i z^C(U_i)$ as $n \to \infty$ will be the same whether outliers always occur consecutively or not. Therefore, we conclude that $P_n^{C|B} = P_n^{C|C}$. From (Eq. 3.1.80), we have

$$P_n^{C|B} = \Phi\left(\frac{m_{f^C} - m_{f_0^{*C}} - v_{f_0^{*C}}\Phi^{-1}(1-\alpha)}{v_{f^C}}\right) \qquad \text{(Eq. 3.1.91)}$$

where $m_{f_0^{*C}}$, $v_{f_0^{*C}}$, m_{f^C}, and v_{f^C} are defined in (Eq. 3.1.65),(Eq. 3.1.66),(Eq. 3.1.78) and (Eq. 3.1.79) respectively. Figure 3-3 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon = 0.1, \mu = -5$. Figure 3-4 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon = 0.01, \mu = -5$. It can easily be in both Figure 3-3 and Figure 3-4 that $P_n^{B|B} > P_n^{A|B}$ and $P_n^{B|B} > P_n^{C|B}$ for all n values plotted. In other words, Figure 3-3 and Figure 3-4 show that the robust detection scheme based on the appropriate outlier model outperforms the two other detection schemes examined.

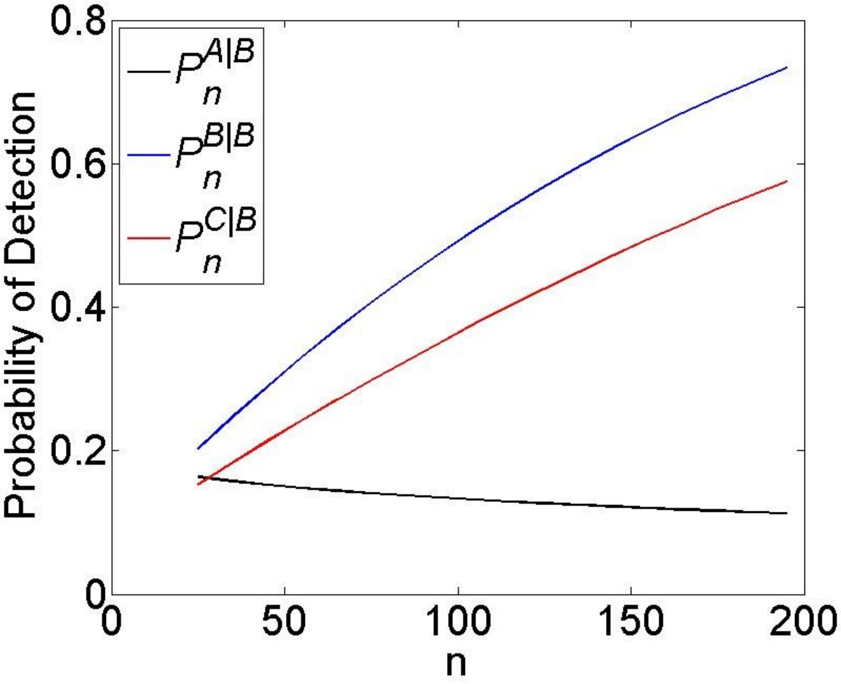

Figure 3-3: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.1, \alpha = 0.05, \mu = -5$.

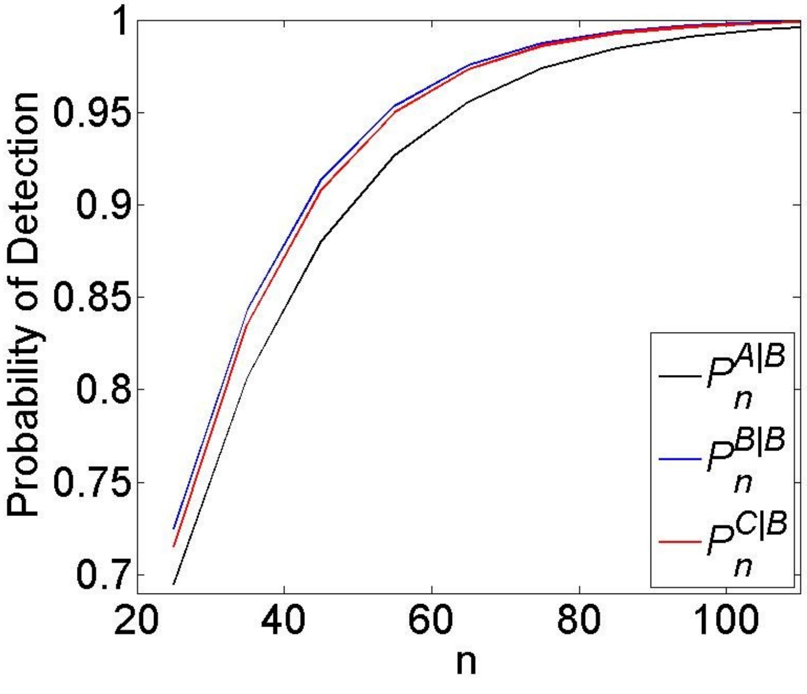

Figure 3-4: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.01, \alpha = 0.05, \mu = -5$.

Comparison under the model of contamination per datum

To find $P_n^{A|C}$, we need to examine the statistics of $n^{-1}\sum_{i=1}^{n} K_i$ where $k_i = x_i\cos\big(\theta_o + p(i)\big) + y_i\sin\big(\theta_o + p(i)\big)$ under

$$f_i(x_i, y_i) = (1-\varepsilon)^2 f_{1,i}(x_i, y_i) + \varepsilon(1-\varepsilon)\big[f_{1,i}(x_i)g_{m,i}(y_i) + g_{m,i}(x_i)f_{1,i}(y_i)\big] + \varepsilon^2 g_{m,i}(x_i)g_{m,i}(y_i) \qquad \text{(Eq. 3.1.92)}$$

Define

$$q_i(k_i): Probability\ density\ of\ k_i\ |f_{1,i}(x_i)g_{m,i}(y_i) \qquad \text{(Eq. 3.1.93)}$$

$$r_i(k_i): Probability\ density\ of\ k_i\ |g_{m,i}(x_i)f_{1,i}(y_i) \qquad \text{(Eq. 3.1.94)}$$

$$s_i(k_i): Probability\ density\ of\ k_i\ |g_{m,i}(x_i)g_{m,i}(y_i) \qquad \text{(Eq. 3.1.95)}$$

Define m_j^A,v_j^A as the mean and standard deviation of $n^{-1}\sum_{i=1}^{n} K_i$ as $n \to \infty$, conditioned on density *j* where *j* can be q, r, s , denoting respectively the densities in (Eq. 3.1.93), (Eq. 3.1.94), and (Eq. 3.1.95). We note that $n^{-1}\sum_{i=1}^{n} K_i$, conditioned on q, r or s, is a linear combination of Gaussian variables and is thus Gaussian. From Section 3.1.2, we note that $g_{m,i}(u_i)$ is Gaussian with mean $mcos(\theta_o + p(i))$; for x_i data values and with mean $msin(\theta_o + p(i))$; for y_i data values, while, both x_i and y_i data values have common variance N_o. Denote $\mu = \max\big(\mu(h)\big)$, where $\mu(h)$ is defined in Section 3.1.2. We then find,

$$m_q^A = E\left\{n^{-1}\sum_{i=1}^{n} K_i \mid q_i\right\} \qquad \text{(Eq. 3.1.96)}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} E\left(x_i|f_{1,i}\right)\cos(\theta) + E\left(y_i|g_m(y_i)\right)\sin(\theta)\, d\theta$$

$$= 2^{-1}\sqrt{E} + 2^{-1}\mu$$

In a similar fashion, we find,

$$v_q^A = v_r^A = v_s^A = \sqrt{n^{-1}N_o} \qquad \text{(Eq. 3.1.97)}$$

$$m_r^A = 2^{-1}\sqrt{E} + 2^{-1}\mu \qquad \text{(Eq. 3.1.98)}$$

$$m_s^A = \mu \qquad \text{(Eq. 3.1.99)}$$

Define $m_{f^{A|C}}$, and $v_{f^{A|C}}$ as the mean and variance of $n^{-1}\sum_{i=1}^{n} K_i$; as $n \to \infty$, conditioned on the density in (Eq. 3.1.92). We then find,

$$m_{f^{A|C}} = (1-\varepsilon)\sqrt{E} + \varepsilon\mu \qquad \text{(Eq. 3.1.100)}$$

$$v_{f^{A|C}}^2 = n^{-1}\left[N_o + \left(\frac{\varepsilon^4}{2} - \varepsilon^3 + \frac{\varepsilon^2}{4} + \frac{\varepsilon}{4}\right)\left(\sqrt{E}-\mu\right)^2\right] \qquad \text{(Eq. 3.1.101)}$$

$$P_n^{A|C} = \Phi\left(\frac{m_{f^{A|C}} - \sqrt{n^{-1}N_o}\,\Phi^{-1}(1-\alpha)}{v_{f^{A|C}}}\right) \qquad \text{(Eq. 3.1.102)}$$

Denote by m_j^B,v_j^B as the mean and standard deviation of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$; as n gets large, conditioned on density j ;where j can be q, r, s, referring respectively to the densities in (Eq. 3.1.93), (Eq. 3.1.94) and (Eq. 3.1.95). We note that k_i is Gaussian when conditioned on q, r or s, with variance N_o. We will define $\mu_j(\theta^*)$ to be the mean of $k(\theta^*)$ divided by $\sqrt{N_o}$, when conditioned on density j ; where j can be q, r or s. We find

$$m_j^B = \frac{1}{2\pi}\int_0^{2\pi} E\{z(K(\theta^*)|j)\}\,d\theta^* \qquad \text{(Eq. 3.1.103)}$$

$$\left(v_j^B\right)^2 = (n)^{-1}\frac{1}{2\pi}\int_0^{2\pi} Var\{z(K(\theta^*)|j)\}\,d\theta^* \qquad \text{(Eq. 3.1.104)}$$

$$\begin{aligned} E\{z(K(\theta^*)|j)\} = \sqrt{N_o}\,\Big\{&\mu_j(\theta^*) \\ &+ \left[\rho - \mu_j(\theta^*) - c(\varepsilon)\right]\Phi\left(c(\varepsilon) - \rho + \mu_j(\theta^*)\right) \\ &+ \left[c(\varepsilon) - \mu_j(\theta^*)\right]\Phi\left(c(\varepsilon) - \mu_j(\theta^*)\right) \\ &+ \phi\left(c(\varepsilon) - \mu_j(\theta^*)\right) - \phi\left(\rho - \mu_j(\theta^*) - c(\varepsilon)\right)\Big\} \end{aligned} \qquad \text{(Eq. 3.1.105)}$$

$$\begin{aligned} &Var\{z(K(\theta^*)|j)\} \\ &\quad = N_o\big\{1 + \mu_j^2(\theta^*) \\ &\quad + \left[c^2(\varepsilon) - \mu_j^2(\theta^*) - 1\right]\Phi\left[c(\varepsilon) - \mu_j(\theta^*)\right] \\ &\quad + \left[(c(\varepsilon) - \rho)^2 - \mu_j^2(\theta^*) - 1\right]\Phi\left[c(\varepsilon) - \rho + \mu_j(\theta^*)\right] \\ &\quad + \left[c(\varepsilon) + \mu_j(\theta^*)\right]\phi\left[c(\varepsilon) - \mu_j(\theta^*)\right] \\ &\quad + \left[c(\varepsilon) - \rho - \mu_j(\theta^*)\right]\phi\left[\rho - \mu_j(\theta^*) - c(\varepsilon)\right] \\ &\quad - N_o^{-1}[E\{z(K(\theta^*)|j)\}]^2\big\} \end{aligned} \qquad \text{(Eq. 3.1.106)}$$

where

$$\mu_q(\theta^*) = N_o^{-1/2}\left[\sqrt{E}\cos^2(\theta^*) + m\sin^2(\theta^*)\right] \qquad \text{(Eq. 3.1.107)}$$

$$\mu_r(\theta^*) = N_o^{-1/2}\left[\sqrt{E}\sin^2(\theta^*) + m\cos^2(\theta^*)\right] \qquad \text{(Eq. 3.1.108)}$$

$$\mu_s(\theta^*) = N_o^{-1/2}m[\cos^2(\theta^*) + \sin^2(\theta^*)] = \mu \qquad \text{(Eq. 3.1.109)}$$

and where (Eq. 3.1.105) and (Eq. 3.1.106) are obtained from (Eq. 3.1.29) and (Eq. 3.1.30). For m_s^B and $(v_s^B)^2$, the dependence of θ^* disappears, and m_s^B and $(v_s^B)^2$ are identical to $m_{g_m}$and $v_{g_m}^2$, which are found in (Eq. 3.1.29) and (Eq. 3.1.30) . Define $m_{f^{B|C}}$ and $v_{f^{B|C}}$ as the mean and variance of $n^{-1}\sum_{i=1}^{n} z(K_i)$;as $n \to \infty$, conditioned on the density in (Eq. 3.1.92). We then find

$$m_{f^{B|C}} = (1-\varepsilon)^2 m_{f_1^B} + \varepsilon(1-\varepsilon) m_q^B + \varepsilon(1-\varepsilon) m_r^B + \varepsilon^2 m_s^B \qquad \text{(Eq. 3.1.110)}$$

$$\begin{aligned} v_{f^{B|C}}^2 = {} & (1-\varepsilon)^2 v_{f_1^B}^2 + \varepsilon(1-\varepsilon)\left[v_q^B\right]^2 + \varepsilon(1-\varepsilon)[v_r^B]^2 + \varepsilon^2 [v_s^B]^2 + \\ & n^{-1}(1-\varepsilon)^2 \left\{ \varepsilon(1-\varepsilon)\left(\left[m_{f_1^B} - m_q^B\right]^2 + \left[m_{f_1^B} - m_r^B\right]^2\right)\right. \\ & \left. + \varepsilon^2 \left[m_{f_1^B} - m_s^B\right]^2 \right\} + \qquad \text{(Eq. 3.1.111)} \\ & n^{-1}\varepsilon^2 \left\{\varepsilon(1-\varepsilon)\left([m_s^B - m_r^B]^2 + \left[m_s^B - m_q^B\right]^2\right)\right\} \\ & + n^{-1}\varepsilon^2(1-\varepsilon)^2\left[m_r^B - m_q^B\right]^2 \end{aligned}$$

and

$$P_n^{B|C} = \Phi\left(\frac{m_{f^{B|C}} - m_{f_0^{*B}} - v_{f_o^{*B}} \Phi^{-1}(1-\alpha)}{v_{f^{B|C}}}\right) \qquad \text{(Eq. 3.1.112)}$$

where $m_{f_0^{*B}}$, and $v_{f_o^{*B}}$ are defined in (Eq. 3.1.23) and (Eq. 3.1.24). Finally, from (Eq. 3.1.80) we have

$$P_n^{C|C} = \Phi\left(\frac{m_{f^C} - m_{f_0^{*C}} - v_{f_o^{*C}} \Phi^{-1}(1-\alpha)}{v_{f^C}}\right) \qquad \text{(Eq. 3.1.113)}$$

where $m_{f_0^{*C}}$, $v_{f_o^{*C}}$, m_{f^C}, and v_{f^C} are defined in (Eq. 3.1.65),(Eq. 3.1.66),(Eq. 3.1.78) and (Eq. 3.1.79), respectively. Figure 3-5 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon =$

$0.1, \mu = -5$. Figure 3-6 shows the power curves for parameters: $E = 0.25, N_o = 1, \varepsilon = 0.01, \mu = -5$. It can easily be seen in both Figure 3-5 and Figure 3-6 that $P_n^{C|C} > P_n^{A|C}$ and $P_n^{C|C} > P_n^{B|C}$ for all n values plotted. In other words, Figure 3-5 and Figure 3-6 show that the robust detection scheme based on the appropriate outlier model outperforms the two other detection schemes examined.

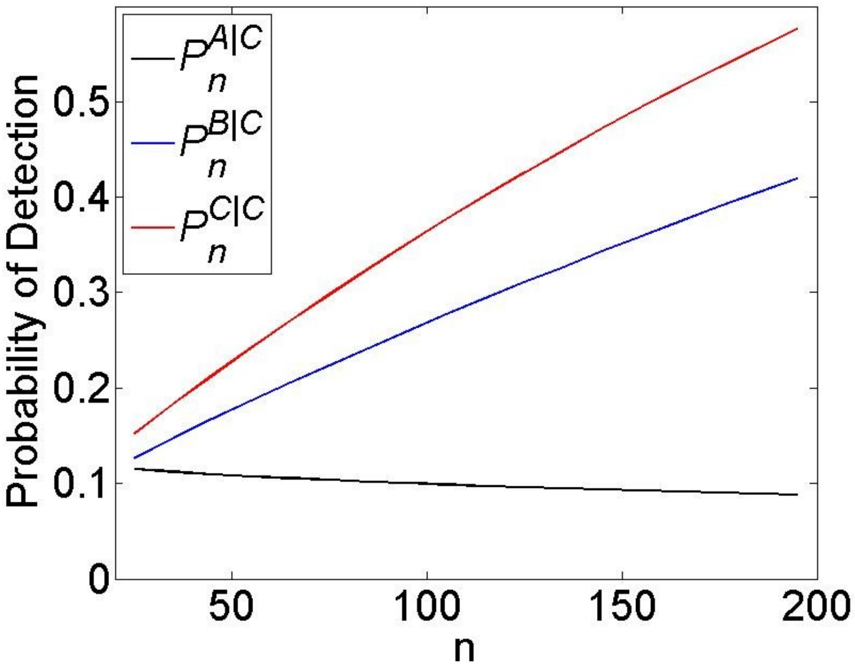

Figure 3-5: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.1, \alpha = 0.05, \mu = -5$.

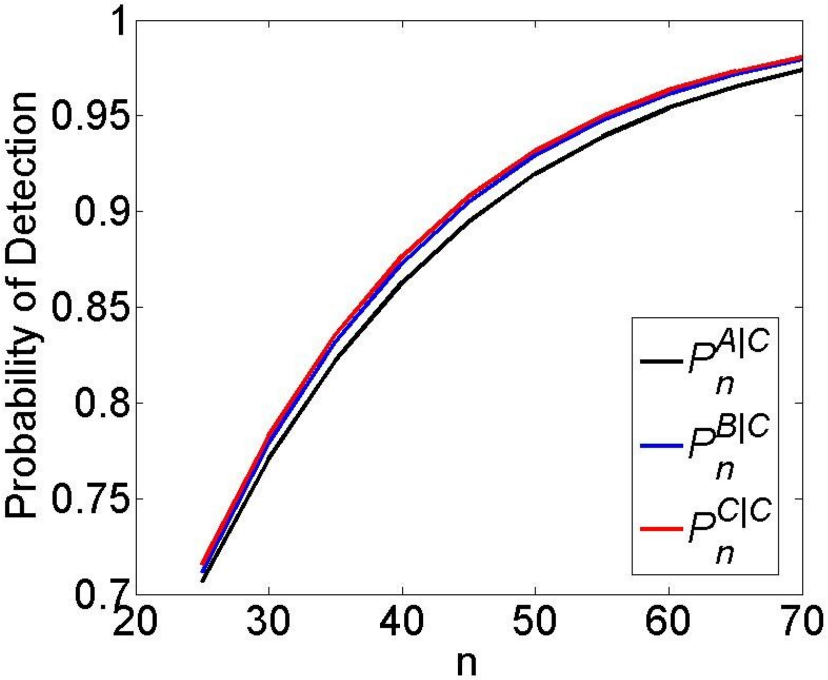

Figure 3-6: Power curves for $E = 0.25, N_o = 1, \varepsilon = 0.01, \alpha = 0.05, \mu = -5$.

3.2 Detection when the waveform is unknown (energy detection)

Let $\mathfrak{I}_1$ and $\mathfrak{I}_o$ be two disjoint classes of n-dimensional density functions, and let $\mathfrak{I}_1$ and $\mathfrak{I}_o$ represent respectively hypotheses H_1 and H_o. Assume that x^n, y^n are generated by a single density function in $\mathfrak{I}_1 \cup \mathfrak{I}_o$. Furthermore, consider that the two classes are Huber classes defined as

$$\mathfrak{I}_o = \{f(x_i, y_i) = (1 - \varepsilon_o) f_o(x_i, y_i) + \varepsilon_o h(x_i, y_i) \qquad \text{(Eq. 3.2.1)}$$

$$\mathfrak{I}_1 = \{f(x_i, y_i) = (1 - \varepsilon_1) f_{1,o}(x_i, y_i) + \varepsilon_1 h(x_i, y_i) \qquad \text{(Eq. 3.2.2)}$$

where

$$f_o(x_i, y_i) = \frac{1}{2\pi N_o} \exp\left(-\frac{1}{2N_o}\left(x_i^2 + y_i^2\right)\right) \qquad \text{(Eq. 3.2.3)}$$

$$f_1(x_i, y_i) = \frac{1}{2\pi N_o} \exp\left(-\frac{1}{2N_o}(x_i^2 + y_i^2) - \frac{E}{2N_o}\right) I_o\left(\frac{\sqrt{x_i^2 E + y_i^2 E}}{N_o}\right) \qquad \text{(Eq. 3.2.4)}$$

and where $\varepsilon_o, \varepsilon_1$ are two given real constants whose values lie in $(0,1)$. The classes above model the cases where a sequence of data is generated by a corresponding nominal process, and contaminations may appear independently per data pair (x_i, y_i) with probabilities ε_o ; for class $\mathfrak{I}_o$ and ε_1 ; for class $\mathfrak{I}_1$.

$$\gamma(x_i, y_i) = \ln\left(\frac{f_1(x_i, y_i)}{f_0(x_i, y_i)}\right) = -\frac{E}{2N_o} + \ln\left(I_o\left(\frac{\sqrt{x_i^2 E + y_i^2 E}}{N_o}\right)\right) \qquad \text{(Eq. 3.2.5)}$$

define $u_i = x_i^2 + y_i^2$. Then

$$\gamma^*(u_i) = \log\left(\frac{f_1^*(u_i)}{f_0^*(u_i)}\right) \qquad \text{(Eq. 3.2.6)}$$

$$= \begin{cases} \ln c_1 + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & u_i\colon \beta(u_i) \le \ln c_1 \\ \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \beta(u_i) & u_i\colon \ln c_1 < \beta(u_i) < \ln c_o \\ \ln c_0 + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & u_i\colon \beta(u_i) \ge \ln c_o \end{cases} \qquad \text{(Eq. 3.2.7)}$$

where

$$\beta(u_i) = -\frac{E}{2N_o} + \ln\left(I_o\left(\frac{\sqrt{u_i E}}{N_o}\right)\right) \qquad \text{(Eq. 3.2.8)}$$

Let us define

$$z_i = I_o\left(\frac{\sqrt{u_i E}}{N_o}\right) \qquad \text{(Eq. 3.2.9)}$$

$$d_1 = c_1 \exp\left(\frac{E}{2N_o}\right) \qquad \text{(Eq. 3.2.10)}$$

$$d_0 = c_o \exp\left(\frac{E}{2N_o}\right) \qquad \text{(Eq. 3.2.11)}$$

Then

$$\gamma^*(z_i) = \log\left(\frac{f_1^*(z_i)}{f_0^*(z_i)}\right) \qquad \text{(Eq. 3.2.12)}$$

$$= \begin{cases} \ln c_1 + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & z_i : z_i \le d_1 \\ \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) - \frac{E}{2N_o} + \ln(z_i) & z_i : d_1 < z_i < d_o \\ \ln c_0 + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) & z_i : z_i \ge d_o \end{cases} \qquad \text{(Eq. 3.2.13)}$$

Finally, define

$$\alpha(z_i) = \gamma^*(z_i) - \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) + \frac{E}{2N_o} \qquad \text{(Eq. 3.2.14)}$$

Then

$$\alpha(z_i) = \begin{cases} \ln d_1 & z_i : z_i \le d_1 \\ \ln z_i & z_i : d_1 < z_i < d_o \\ \ln d_o & z_i : z_i \ge d_o \end{cases} \qquad \text{(Eq. 3.2.15)}$$

$$T^*(u^n) = n^{-1} \sum_{i=1}^{n} \left[\alpha(u_i) + \ln\left(\frac{1-\varepsilon_1}{1-\varepsilon_0}\right) - \frac{E}{2N_o}\right] \qquad \text{(Eq. 3.2.16)}$$

The robust detection rule can be written as

$$\delta^*(u^n) = \begin{cases} 1 \quad u^n\colon n^{-1}\sum_{i=1}^{n}\alpha(u_i) > \lambda^* \\ r^* \quad u^n\colon n^{-1}\sum_{i=1}^{n}\alpha(u_i) = \lambda^* \\ 0 \quad u^n\colon n^{-1}\sum_{i=1}^{n}\alpha(u_i) < \lambda^* \end{cases} \qquad \text{(Eq. 3.2.17)}$$

where λ^* and r^* are such that

$$\Pr\left\{n^{-1}\sum_{i=1}^{n}\alpha(u_i) > \lambda^* \,\middle|\, f_0^*\right\} + r^*\Pr\left\{n^{-1}\sum_{i=1}^{n}\alpha(u_i) = \lambda^* \,\middle|\, f_0^*\right\} = \alpha \qquad \text{(Eq. 3.2.18)}$$

The least favorable density functions f_0^* and f_1^* are

$$f_0^*(u) = \begin{cases} \dfrac{(1-\varepsilon_0)}{2N_o}\exp\left(-\dfrac{u}{2N_o}\right) & u\colon I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) < d_0 \\ \dfrac{(1-\varepsilon_0)}{\mathrm{2d_oN_0}}\exp\left(\dfrac{-u}{2N_o}\right) I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) & u\colon I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) \geq d_0 \end{cases} \qquad \text{(Eq. 3.2.19)}$$

$$f_1^*(u) = \begin{cases} \dfrac{(1-\varepsilon_1)}{2N_o}\exp\left(\dfrac{-u-E}{2N_o}\right) I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) & u\colon I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) > d_1 \\ \dfrac{(1-\varepsilon_1)d_1}{\mathrm{2N_0}}\exp\left(-\dfrac{E}{2N_o}\right)\exp\left(\dfrac{-u}{2N_o}\right) & u\colon I_o\left(\dfrac{\sqrt{uE}}{N_o}\right) \leq d_1 \end{cases} \qquad \text{(Eq. 3.2.20)}$$

and the constants d_0 and d_1 are such that

$$\begin{aligned} &(1-\varepsilon_0)[\Pr\left\{I_o\left(\frac{\sqrt{UE}}{N_o}\right) < d_o \,\middle|\, f_o\right\} \\ &\qquad + \frac{\exp\left(\frac{E}{2N_o}\right)}{\mathrm{d_o}}\Pr\left\{I_o\left(\frac{\sqrt{UE}}{N_o}\right) \geq d_o \,\middle|\, f_1\right\}] = 1 \end{aligned} \qquad \text{(Eq. 3.2.21)}$$

$$(1-\varepsilon_1)[\Pr\left\{I_o\left(\frac{\sqrt{UE}}{N_o}\right) > d_1 \middle| f_1\right\} + \mathrm{d}_1 \exp\left(-\frac{\mathrm{E}}{2\mathrm{N_o}}\right)\Pr\left\{I_o\left(\frac{\sqrt{UE}}{N_o}\right) \le d_1 \middle| f_o\right\}] = 1 \qquad \text{(Eq. 3.2.22)}$$

Define $I_0^{-1}(x)$, such that $I_0^{-1}(x) = y$; if and only if $I_0\ (y) = x$. Define new constants as below,

$$k_0 = \frac{N_0^2}{E}(I_0^{-1}(d_o))^2 \qquad \text{(Eq. 3.2.23)}$$

$$k_1 = \frac{N_0^2}{E}(I_0^{-1}(d_1))^2 \qquad \text{(Eq. 3.2.24)}$$

With this notation, (Eq. 3.2.15) can be written

$$\alpha(u_i) = \begin{cases} \ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right) & u_i\colon u_i \le k_1 \\ \ln\left(I_o\left(\frac{\sqrt{u_iE}}{N_o}\right)\right) & u_i\colon k_1 < u_i < k_o \\ \ln\left(I_o\left(\frac{\sqrt{k_oE}}{N_o}\right)\right) & u_i\colon u_i \ge k_o \end{cases} \qquad \text{(Eq. 3.2.25)}$$

Then, (Eq. 3.2.21) and (Eq. 3.2.22) can be written

$$(1-\varepsilon_0)[\Pr\{U < k_o|f_o\} + \frac{\exp\left(\frac{E}{2N_o}\right)}{\mathrm{I_o}\left(\frac{\sqrt{\mathrm{k_oE}}}{\mathrm{N_o}}\right)}\Pr\{U \ge k_o|f_1\}] = 1 \qquad \text{(Eq. 3.2.26)}$$

$$(1-\varepsilon_1)[\Pr\{U > k_1|f_1\} + \mathrm{I_o}\left(\frac{\sqrt{\mathrm{k_1E}}}{\mathrm{N_o}}\right)\exp\left(-\frac{\mathrm{E}}{2\mathrm{N_o}}\right)\Pr\{U \le k_1|f_o\}] = 1 \qquad \text{(Eq. 3.2.27)}$$

or equivalently as; if $k_0, k_1 \ge 0$:

$$(1-\varepsilon_0)[1-\exp\left(-\frac{k_o}{2N_o}\right) \quad + \frac{\exp\left(\frac{E}{2N_o}\right)}{\mathrm{I_o}\left(\sqrt{\frac{\mathrm{k_0 E}}{\mathrm{N_o^2}}}\right)} \mathrm{Q_1}(\sqrt{\mathrm{E/N_o}}, \sqrt{\mathrm{k_0/N_o}}\,)\ ] = 1 \qquad \text{(Eq. 3.2.28)}$$

$$(1-\varepsilon_1)[\mathrm{Q_1}(\sqrt{\mathrm{E/N_o}}\,, \sqrt{\mathrm{k_1/N_0}}) + \mathrm{I_o}\left(\sqrt{\frac{\mathrm{k_1 E}}{\mathrm{N_o^2}}}\right)\exp\left(-\frac{\mathrm{E}}{2\mathrm{N_o}}\right)(1-\exp\left(-\frac{k_1}{2N_o}\right))\ ] = 1 \qquad \text{(Eq. 3.2.29)}$$

We note that if $k_1 < 0$, then (Eq. 3.2.27) becomes $1 - \varepsilon_1 = 1$; therefore, this will only occur if $\varepsilon_1 = 0$. Also we note that we must have $k_1 \leq k_0$ as dictated by (Eq. 3.2.25).

We will focus our analysis on the case where $\varepsilon_o = \varepsilon_1 = \varepsilon$.

We will denote the solutions of k_o, k_1 in (Eq. 3.2.28), (Eq. 3.2.29), as $k_o(E, N_o, \varepsilon)$ and $k_1(E, N_o, \varepsilon)$. $k_o(E, N_o, \varepsilon)$ and $k_1(E, N_o, \varepsilon)$ are continuous since all functions in (Eq. 3.2.28) and (Eq. 3.2.29) are continuous over all variables.

We will define the break-down point $\varepsilon_b(E, N_o)$ as the ε closest to zero such that $k_o(E, N_o, \varepsilon) \leq k_1(E, N_o, \varepsilon)$. We can also denote by k_b the k value at which the breakdown point occurs; in other words the value when $k_o = k_1$ occurs. From (Eq. 3.2.28) and (Eq. 3.2.29) with $k_o = k_1 = k_b$ and $\varepsilon_o = \varepsilon_1 = \varepsilon$, we conclude that k_b must satisfy

$$\mathrm{I_o}\left(\sqrt{\frac{\mathrm{k_b E}}{\mathrm{N_o^2}}}\right) = \exp\left(\frac{\mathrm{E}}{2\mathrm{N_o}}\right) \qquad \text{(Eq. 3.2.30)}$$

and

$$\varepsilon_b = \frac{Q_1(\sqrt{E/N_o}, \sqrt{k_b/N_o}\,) - \exp\left(-\frac{k_b}{2N_o}\right)}{1 + Q_1(\sqrt{E/N_o}, \sqrt{k_b/N_o}\,) - \exp\left(-\frac{k_b}{2N_o}\right)} \qquad \text{(Eq. 3.2.31)}$$

We note that since

$$Q_1(\sqrt{E/N_o}, \sqrt{k_b/N_o}\,) - \exp\left(-\frac{k_b}{2N_o}\right) < 1 \qquad \text{(Eq. 3.2.32)}$$

then

$$\varepsilon_b < 0.5 \qquad \text{(Eq. 3.2.33)}$$

There is no known analytic expression for $k_o(E, N_o, \varepsilon)$, $k_1(E, N_o, \varepsilon)$, $k_b(E, N_o)$ or $\varepsilon_b(E, N_o)$; therefore we will use numerical methods to analyze these equations.

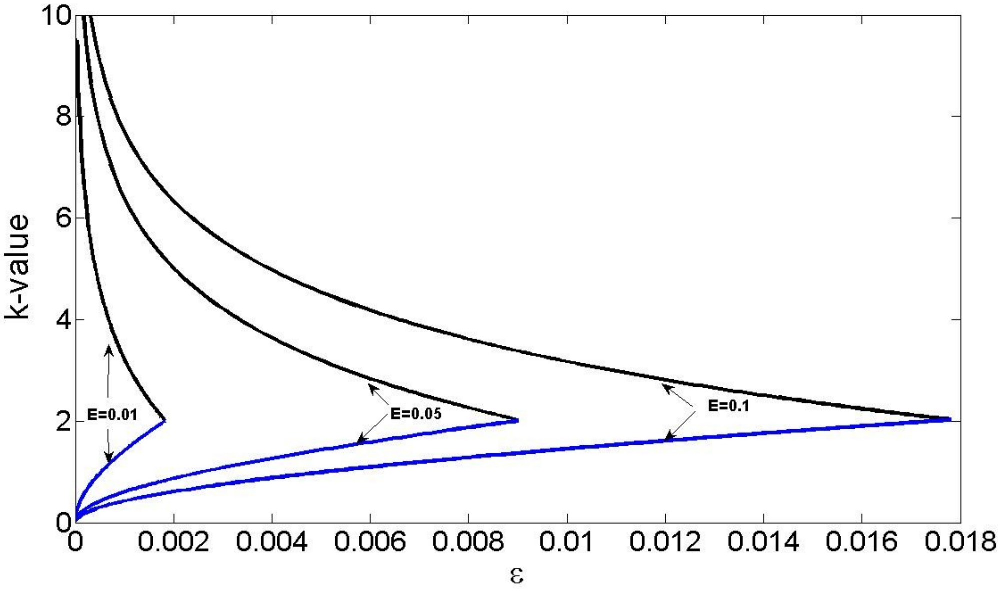

Figure 3-7: Threshold Values when $E < N_o$ ($N_o = 1$).

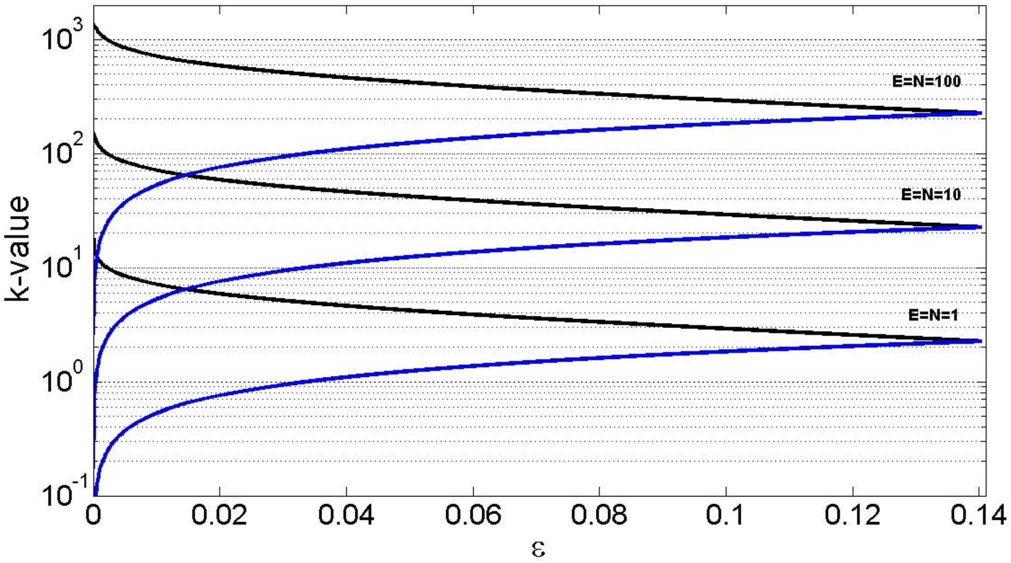

Figure 3-8: Threshold Values when $E = N_o$.

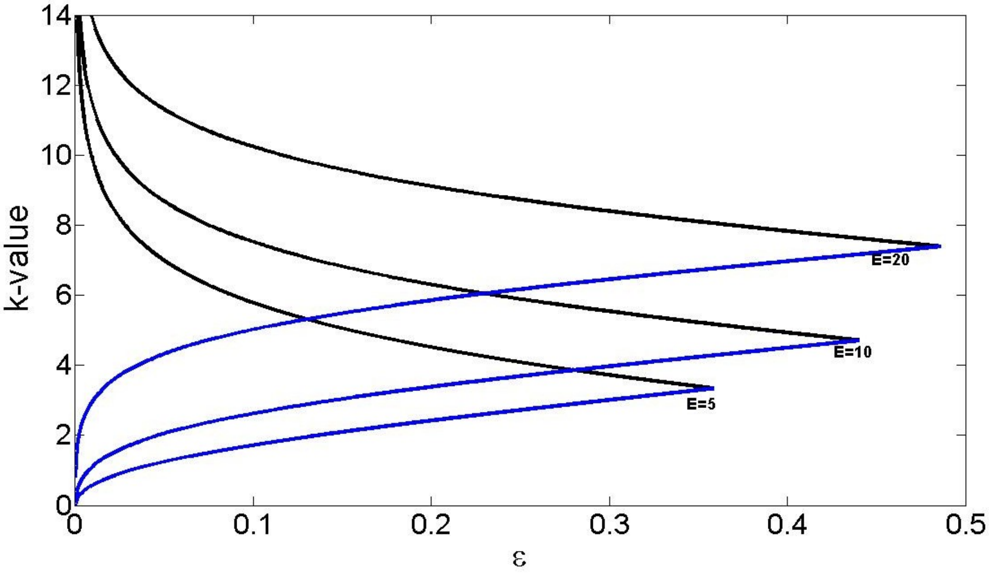

Figure 3-9: Threshold Values when $E > N_o \quad (N_o = 1)$.

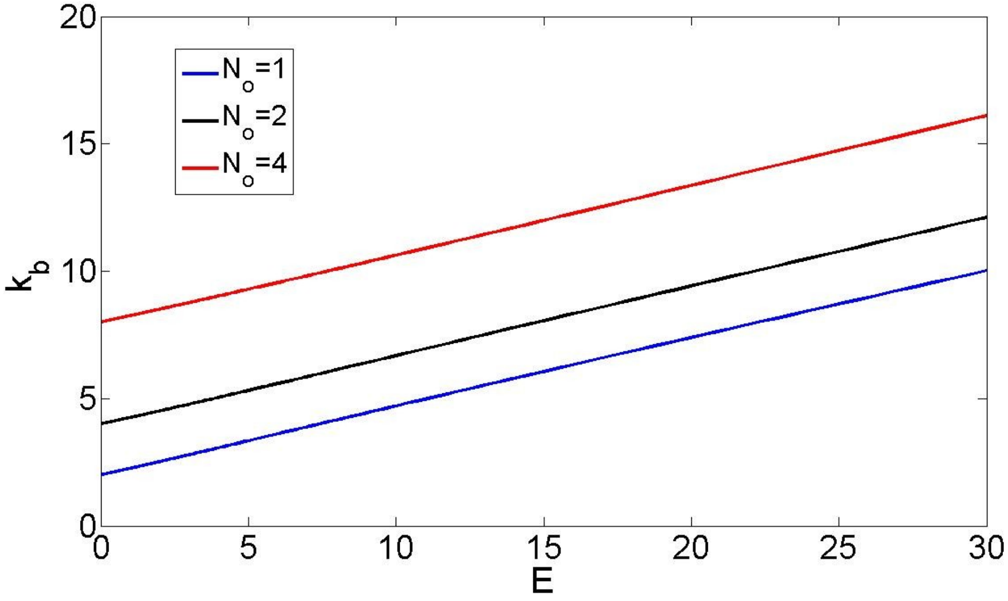

Figure 3-10: Threshold Value at the breakdown point.

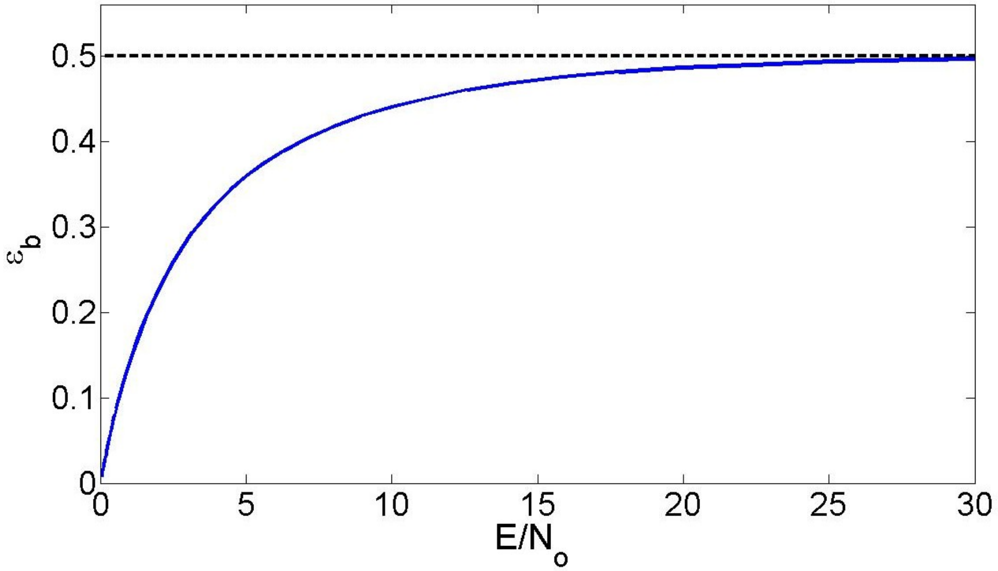

Figure 3-11: Breakdown Point $(N_o = 1)$.

Since the variables $\{U_i\}$ are independent and identically distributed, when conditioned on any f in $\mathfrak{I}_1 \cup \mathfrak{I}_o$, the random variables $\{\beta_i = \alpha(U_i)\}$ are too, and the random variable $n^{-1}\sum_{i=1}^{n}\alpha(U_i)$ is then asymptotically Gaussian. From (Eq. 3.2.21), (Eq. 3.2.22) and (Eq. 3.2.25) we find

$$m_{f_0^*} = E\{n^{-1}\sum_{i=1}^{n}\alpha(U)|f_o^*\} = E\{\alpha(U)|f_o^*\} \qquad \text{(Eq. 3.2.34)}$$

$$\begin{aligned} & m_{f_0^*} \\ & = \int_0^{k_1} \ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right)\frac{(1-\varepsilon)}{2N_o}\exp\left(-\frac{u}{2N_o}\right)du \\ & + \int_{k_1}^{k_o} \ln\left(I_o\left(\frac{\sqrt{uE}}{N_o}\right)\right)\frac{(1-\varepsilon)}{2N_o}\exp\left(-\frac{u}{2N_o}\right)du \qquad \text{(Eq. 3.2.35)} \\ & + \left(I_o\left(\frac{\sqrt{k_oE}}{N_o}\right)\right)^{-1}\int_{k_o}^{\infty} \ln\left(I_o\left(\frac{\sqrt{Ek_0}}{N_o}\right)\right)\frac{(1-\varepsilon)}{2N_o}\exp\left(-\frac{u}{2N_o}\right)I_o\left(\frac{\sqrt{uE}}{N_o}\right)du \end{aligned}$$

$$\begin{aligned} m_{f_0^*} &= (1-\varepsilon)\left(1-\exp\left(-\frac{k_1}{2N_o}\right)\right)\ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right) \\ & + \int_{k_1}^{k_o} \ln\left(I_o\left(\frac{\sqrt{uE}}{N_o}\right)\right)\frac{(1-\varepsilon)}{2N_o}\exp\left(-\frac{u}{2N_o}\right)du \qquad \text{(Eq. 3.2.36)} \\ & +(1-\varepsilon)\ln\left(I_o\left(\frac{\sqrt{k_0E}}{N_o}\right)\right)\left(I_o\left(\frac{\sqrt{k_oE}}{N_o}\right)\right)^{-1}\exp\left(\frac{E}{2N_o}\right)Q_1\left(\sqrt{\frac{E}{N_o}},\sqrt{\frac{k_o}{N_o}}\right) \end{aligned}$$

The integral in (Eq. 3.2.36) cannot be written in terms of a well-known function. In the spirit of the Marcum Q function we define the following functions

$$R_1(a,b,c) = \int_b^c xln\big(I_o(ax)\big)\exp\left(-\frac{x^2}{2}-\frac{a^2}{2}\right)dx \qquad \text{(Eq. 3.2.37)}$$

$$R_2(a,b,c) = \int_b^c x\left(ln\big(I_o(ax)\big)\right)^2\exp\left(-\frac{x^2}{2}-\frac{a^2}{2}\right)dx \qquad \text{(Eq. 3.2.38)}$$

Then (Eq. 3.2.36) can be written as,

$$
\begin{aligned}
m_{f_0^*} &= (1-\varepsilon)\left(1-\exp\left(-\frac{k_1}{2N_o}\right)\right)\ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right) \\
&+ (1-\varepsilon)\exp\left(\frac{E}{2N_o}\right)R_1\left(\sqrt{\frac{E}{N_o}},\sqrt{\frac{k_1}{N_o}},\sqrt{\frac{k_o}{N_o}}\right) \\
&+(1-\varepsilon)\ln\left(I_o\left(\frac{\sqrt{k_0E}}{N_o}\right)\right)\left(I_o\left(\frac{\sqrt{k_oE}}{N_o}\right)\right)^{-1}\exp\left(\frac{E}{2N_o}\right)Q_1\left(\sqrt{\frac{E}{N_o}},\sqrt{\frac{k_o}{N_o}}\right)
\end{aligned}
\qquad \text{(Eq. 3.2.39)}
$$

and

$$
v_{f_0^*}^2 = Var\{n^{-1}\sum_{i=1}^{n}\alpha(U)|f_o^*\} = n^{-1}Var\{\alpha(U)|f_o^*\} \qquad \text{(Eq. 3.2.40)}
$$

$$
\begin{aligned}
&= n^{-1}(1-\varepsilon)\left(1-\exp\left(-\frac{k_1}{2N_o}\right)\right)\left(ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right)\right)^2 \\
&+ n^{-1}(1-\varepsilon)\exp\left(\frac{E}{2N_o}\right)R_2\left(\sqrt{\frac{E}{N_o}},\sqrt{\frac{k_1}{N_o}},\sqrt{\frac{k_o}{N_o}}\right) \\
&+n^{-1}(1 \\
&-\varepsilon)\left(\ln\left(I_o\left(\frac{\sqrt{k_0E}}{N_o}\right)\right)\right)^2\left(I_o\left(\frac{\sqrt{k_oE}}{N_o}\right)\right)^{-1}\exp\left(\frac{E}{2N_o}\right)Q_1\left(\sqrt{\frac{E}{N_o}},\sqrt{\frac{k_o}{N_o}}\right) \\
&-n^{-1}m_{f_0^*}^2
\end{aligned}
\qquad \text{(Eq. 3.2.41)}
$$

For asymptotically large values of n, the constant r^* in (Eq. 3.2.18) can be set at zero, resulting in,

$$\lambda^* = m_{f_0^*} + v_{f_0^*}\, \Phi^{-1}(1-\alpha). \qquad \text{(Eq. 3.2.42)}$$

We will examine the asymptotic performance of our detector under $f = (1-\varepsilon)f_1 + \varepsilon g_m;$ where

$$f_1(u) = \frac{1}{2N_o} \exp\left(-\frac{u}{2N_o} - \frac{E}{2N_o}\right) I_o\left(\frac{\sqrt{uE}}{N_o}\right) \qquad \text{(Eq. 3.2.43)}$$

and

$$g_m(u) = \frac{1}{2N_{g_m}} \exp\left(-\frac{u}{2N_{g_m}} - \frac{\gamma}{2N_{g_m}}\right) I_o\left(\frac{\sqrt{u\gamma}}{N_{g_m}}\right) \qquad \text{(Eq. 3.2.44)}$$

and where γ is a parameter of $g_m(u)$, which may be viewed as the energy of the outliers. We note that $f_1(u)$ is the same density as that given by (Eq. 3.2.4), written here in terms of u instead of (x, y). Also, $g_m(u)$ results from assuming a Gaussian model outlier under both x and y; with means m_x and m_y , such that $\gamma = m_x^2 + m_y^2$ and variance summing to $2N_{g_m}$. Let m_{f_1}, $v_{f_1}^2$, m_{g_m}, and $v_{g_m}^2$ denote the mean and variance of $n^{-1}\sum_{i=1}^{n} \alpha(U_i)$ conditioned on f_1 or g_m , respectively. We define the following useful functions

$$R_3(a,b,c,d) = \int_a^b x ln\big(I_o(ax)\big) \exp\left(-\frac{x^2}{2} - \frac{b^2}{2}\right) I_o(bx)dx \qquad \text{(Eq. 3.2.45)}$$

$$R_4(a,b,c,d) = \int_c^d x\left(ln\big(I_o(ax)\big)\right)^2 \exp\left(-\frac{x^2}{2} - \frac{b^2}{2}\right) I_o(bx)dx \qquad \text{(Eq. 3.2.46)}$$

We obtain

$$\begin{aligned} m_{f_1} = {} & \ln\left(I_o\left(\frac{\sqrt{k_1 E}}{N_o}\right)\right)\left(1 - Q_1\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}\right)\right) \\ & + \ln\left(I_o\left(\frac{\sqrt{k_0 E}}{N_o}\right)\right) Q_1\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_0}{N_o}}\right) \\ & + R_3\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}, \sqrt{\frac{k_o}{N_o}}\right) \end{aligned} \tag{Eq. 3.2.47}$$

$$\begin{aligned} v_{f_1}^2 = {} & n^{-1}\left[\ln\left(I_o\left(\frac{\sqrt{k_1 E}}{N_o}\right)\right)\right]^2 \left(1 - Q_1\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}\right)\right) \\ & + n^{-1}\left[\ln\left(I_o\left(\frac{\sqrt{k_0 E}}{N_o}\right)\right)\right]^2 Q_1\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_0}{N_o}}\right) \\ & + n^{-1} R_4\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}, \sqrt{\frac{k_o}{N_o}}\right) - n^{-1} m_{f_1}^2 \end{aligned} \tag{Eq. 3.2.48}$$

$$\begin{aligned} m_{g_m} = {} & \ln\left(I_o\left(\frac{\sqrt{k_1 E}}{N_o}\right)\right)\left(1 - Q_1\left(\sqrt{\frac{\gamma}{N_{g_m}}}, \sqrt{\frac{k_1}{N_{g_m}}}\right)\right) \\ & + \ln\left(I_o\left(\frac{\sqrt{k_0 E}}{N_o}\right)\right) Q_1\left(\sqrt{\frac{\gamma}{N_{g_m}}}, \sqrt{\frac{k_0}{N_{g_m}}}\right) \\ & + R_3\left(\frac{\sqrt{E N_{g_m}}}{N_0}, \sqrt{\frac{\gamma}{N_{g_m}}}, \sqrt{\frac{k_1}{N_{g_m}}}, \sqrt{\frac{k_o}{N_{g_m}}}\right) \end{aligned} \tag{Eq. 3.2.49}$$

$$\begin{aligned} v_{g_m}^2 = n^{-1}&\left[\ln\left(I_o\left(\frac{\sqrt{k_1E}}{N_o}\right)\right)\right]^2\left(1-Q_1\left(\sqrt{\frac{\gamma}{N_{g_m}}},\sqrt{\frac{k_1}{N_{g_m}}}\right)\right)\\ &+n^{-1}\left[\ln\left(I_o\left(\frac{\sqrt{k_0E}}{N_o}\right)\right)\right]^2 Q_1\left(\sqrt{\frac{\gamma}{N_{g_m}}},\sqrt{\frac{k_0}{N_{g_m}}}\right)\\ &+n^{-1}R_4\left(\frac{\sqrt{EN_{g_m}}}{N_0},\sqrt{\frac{\gamma}{N_{g_m}}},\sqrt{\frac{k_1}{N_{g_m}}},\sqrt{\frac{k_o}{N_{g_m}}}\right)\\ &-n^{-1}m_{g_m}^2 \end{aligned} \tag{Eq. 3.2.50}$$

For

$$f=(1-\varepsilon)f_1+\varepsilon g_m \tag{Eq. 3.2.51}$$

we have

$$m_f=(1-\varepsilon)m_{f_1}+\varepsilon m_{g_m} \tag{Eq. 3.2.52}$$

$$v_f^2=(1-\varepsilon)v_{f_1}^2+\varepsilon v_{g_m}^2+n^{-1}\varepsilon(1-\varepsilon)\left[m_{f_1}-m_{g_m}\right]^2 \tag{Eq. 3.2.53}$$

Let us denote by $P_n(f,f_o^*)$ the power induced by the robust decision rule in (Eq. 3.2.18) at the density function f in (Eq. 3.2.51). Due to (Eq. 3.2.42), we conclude; for $n\to\infty$,

$$P_n(f,f_o^*)=\Phi\left(\frac{m_f-m_{f_0^*}-v_{f_o^*}\Phi^{-1}(1-\alpha)}{v_f}\right) \tag{Eq. 3.2.54}$$

Given a false alarm rate α, let us consider the optimal Neyman-Pearson test at the density functions f_1 and f_o. By (Eq. 2.3.10) we have

$$P_n^0(f_1, f_o) = Q_n\left(\sqrt{\frac{nE}{N_o}}, \sqrt{H_n^{-1}(1-\alpha)}\right) \qquad \text{(Eq. 3.2.55)}$$

where

$$H_n(x) = \int_0^x d\tau \frac{1}{2^n(n-1)!}(\tau)^{n-1} \exp\left(-\frac{x}{2}\right) \qquad \text{(Eq. 3.2.56)}$$

and $H_n^{-1}(x)$ is the inverse function of $H_n(x)$. $H_n(x)$ is the cumulative distribution function of the gamma random variable with parameters 2 and n. Denote by $m(f)$ and $v^2(f)$ the mean and variance of U under the density function $f = (1-\varepsilon)f_1 + \varepsilon g_m$. Then,

$$\begin{aligned} m(f) &= (1-\varepsilon)E(u|f_1) + \varepsilon E(u|g_m) \\ &= (1-\varepsilon)(E + 2N_o) + \varepsilon(\gamma + 2N_{g_m}) \end{aligned} \qquad \text{(Eq. 3.2.57)}$$

$$\begin{aligned} v^2(f) = 4(1-\varepsilon)(EN_o + N_o^2) + 4\varepsilon\left(\gamma N_{g_m} + N_{g_m}^2\right) \\ + \varepsilon(1-\varepsilon)\left(E + 2N_o - \gamma - 2N_{g_m}\right)^2 \end{aligned} \qquad \text{(Eq. 3.2.58)}$$

Let us denote by $P_n^0(f, f_o)$ the power induced by the optimal Neyman person rule at f, for $n \to \infty$. Directly from (Eq. 2.3.8), (Eq. 3.2.57), and (Eq. 3.2.58), we obtain

$$P_n^0(f, f_o) = \Phi\left(\frac{m(f)\sqrt{n} - N_o\, H_n^{-1}(1-\alpha)/\sqrt{n}}{v(f)}\right) \qquad \text{(Eq. 3.2.59)}$$

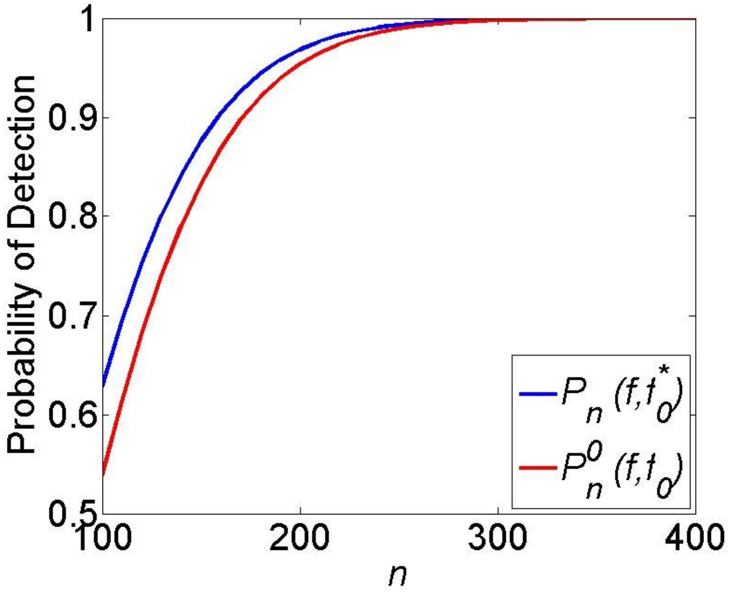

Figure 3-12: Probability of detection with $E = 1, N_o = 1, \alpha = 10^{-5}, \varepsilon = 0.001, \gamma = 0.1$, and $N_{g_m} = 0.1$.

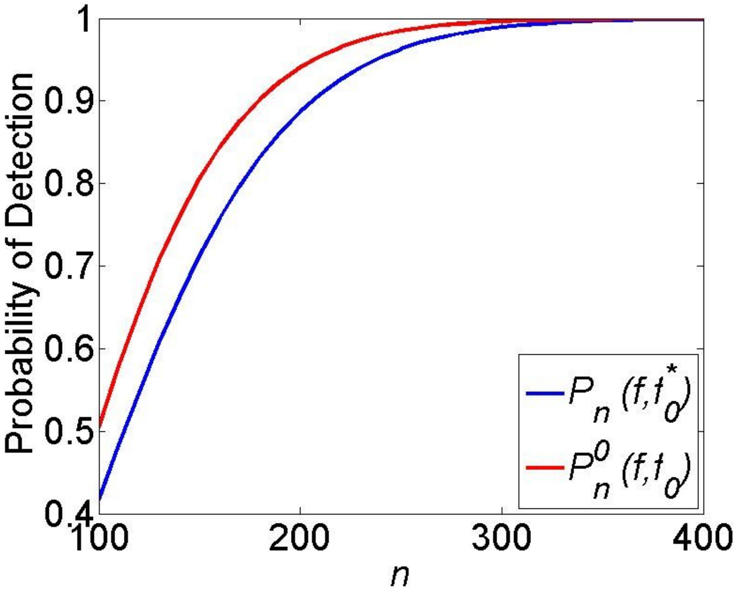

Figure 3-13: Probability of detection with $E = 1, N_o = 1, \alpha = 10^{-5}, \varepsilon = 0.01, \gamma = 0.1$, and $N_{g_m} = 0.1$.

In Figure 3-12, we observe that $P_n(f, f_o^*) > P_n^0(f, f_o)$;for all n examined. However, in Figure 3-13, we observe that $P_n(f, f_o^*) < P_n^0(f, f_o)$; for all n examined. The only difference between Figure 3-12 and Figure 3-13 is that in Figure 3-12, $\varepsilon = 0.001$ while in Figure 3-13 $\varepsilon = 0.01$. The test in $P_n(f, f_o^*)$ set a threshold to guarantee that the probability of false alarm was less than or equal to α at f_o^*, and subsequently guaranteed that the probability of false alarm at any $f \in \mathfrak{I}_0$ was less than or equal to α. In contrast, $P_n^0(f, f_o)$ set a threshold to guarantee that the probability of false alarm was less than or equal to α at f_o only. Let β denote the smallest false alarm satisfied by the test used to produce $P_n(f, f_o^*)$ at f_o. The robust detection guarantees that $\beta \leq \alpha$. If $\beta \ll \alpha$, then comparing $P_n(f, f_o^*)$ and $P_n^0(f, f_o)$ may be misleading, since each satisfies different false alarm upper bounds at f_o. Let us examine $P_n(f, f_o)$, the power of the robust decision rule; when the threshold is selected to satisfy the false alarm constraint only at f_o. We then obtain,

$$P_n(f, f_o^*) = \Phi\left(\frac{m_f - m_{f_o} - v_{f_o}\Phi^{-1}(1-\alpha)}{v_f}\right) \qquad \text{(Eq. 3.2.60)}$$

where

$$m_{f_o} = \left(1 - \exp\left(-\frac{k_1}{2N_o}\right)\right) \ln\left(I_o\left(\frac{\sqrt{k_1 E}}{N_o}\right)\right)$$

$$+ \exp\left(\frac{E}{2N_o}\right) R_1\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}, \sqrt{\frac{k_o}{N_o}}\right) + \ln\left(I_o\left(\frac{\sqrt{k_0 E}}{N_o}\right)\right) \exp\left(-\frac{k_o}{2N_o}\right) \qquad \text{(Eq. 3.2.61)}$$

$$v_{f_0}^2 = n^{-1}\left(1 - \exp\left(-\frac{k_1}{2N_o}\right)\right)\left(ln\left(I_o\left(\frac{\sqrt{k_1 E}}{N_o}\right)\right)\right)^2$$

$$+ n^{-1} \exp\left(\frac{E}{2N_o}\right) R_2\left(\sqrt{\frac{E}{N_o}}, \sqrt{\frac{k_1}{N_o}}, \sqrt{\frac{k_o}{N_o}}\right) \qquad \text{(Eq. 3.2.62)}$$

$$+ n^{-1}\left(\ln\left(I_o\left(\frac{\sqrt{k_0 E}}{N_o}\right)\right)\right)^2 \exp\left(-\frac{k_o}{2N_o}\right) - n^{-1} m_{f_0^*}^2$$

Figure 3-14 shows that $P_n(f, f_0) > P_n^0(f, f_o)$ for all n examined, using the same parameters as in Figure 3-13. At f_1, the optimal test at the nominal model is expected to outperform the robust detection rule. The power loss due to the use of the robust detection rule is controlled by ε, and so is the density used to set the threshold of the test; in satisfaction of the false alarm constraint. Figure 3-15 shows $P_n(f_1, f_0^*)$, $P_n(f_1, f_0)$ and $P_n^0(f_1, f_o)$; with $\varepsilon = 0.05$. It is noted that the power loss in $P_n(f_1, f_0)$ is much smaller than that in $P_n(f_1, f_0^*)$; Figure 3-15. This is evidence of the penalty paid when a false alarm of less than or equal to α is guaranteed at any $\in \mathfrak{I}_0$; instead of only at f_0.

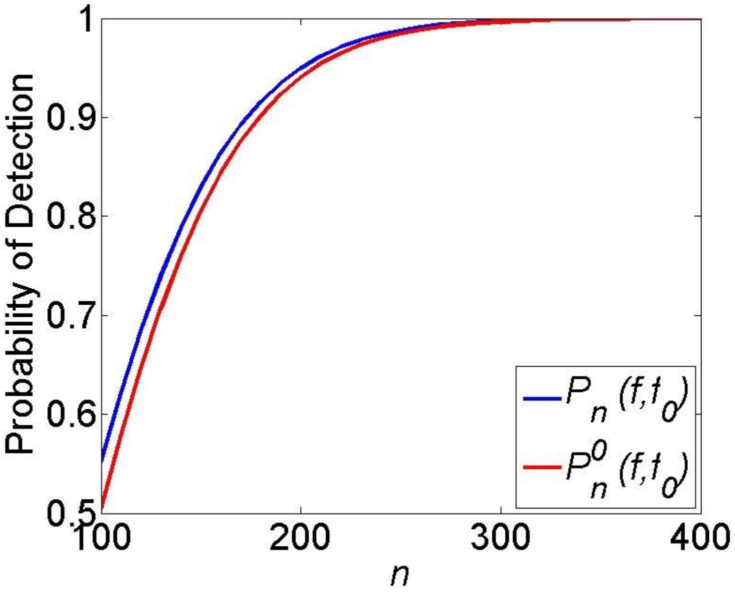

Figure 3-14: Probability of detection with $E = 1, N_o = 1, \alpha = 10^{-5}, \varepsilon = 0.01, \gamma = 0.1$, and $N_{g_m} = 0.1$.

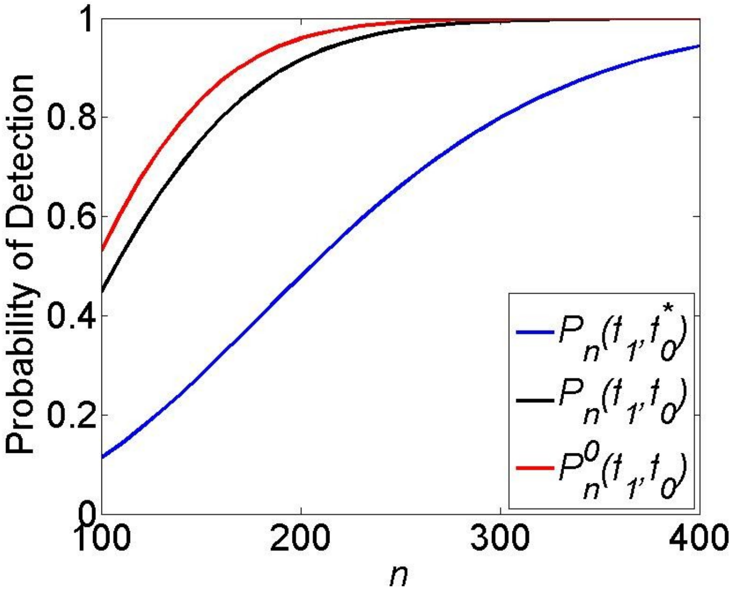

Figure 3-15: Probability of detection with $E = 1, N_o = 1, \alpha = 10^{-5}$ and ε set to 0.05.

3.3 Detection when the waveform is known down to an unknown initial phase

In this case, $f_1(x^n, y^n)$ is not induced by a memoryless process, and the theorems- due to Huber- used in sections 3.1 and 3.2 cannot be used in the search for the optimal robust decision rule. We can, however, use the results from section 3.1 to construct robust decision rules which should provide good protection against occasional outliers. In this section we will develop such ad-hoc, but intuitively pleasing decision rules, and analyze their performance. For more work on robust detection of processes with memory, the reader is referred to [4].

We will begin by deriving a Neyman Pearson like rule for the non robust case. The reason for doing this will become obvious later on. In section 2.4 we derived our decision rule based on the metric $\log \frac{f_{H1}}{f_{HO}}$ where

$$f_{H1}(x^n, y^n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{H1}(x^n, y^n | \boldsymbol{\theta_o} = \theta) d\theta \qquad \text{(Eq. 3.3.1)}$$

In this section, we will examine the Neyman Pearson like rule based on the metric $\log \frac{f_{H1,\theta_{\mathrm{mx}}}}{f_{HO}}$ where

$$f_{H1,\theta_{\mathrm{mx}}} = \max\{f_{H1}(x^n, y^n | \theta) \colon \theta\} \qquad \text{(Eq. 3.3.2)}$$

and θ_{mx} is the θ which maximizes the conditional probability density function. Using the notation in section 2.4, we find

$$f_{H1}(x^n, y^n | \theta) = \frac{1}{(2\pi)^n N_o^n} \exp\left(-\frac{1}{2N_0}[T^2 + nE]\right) \exp\left(\frac{\sqrt{E}}{N_0} V * Re\{\exp(j[\omega + \theta])\}\right) \qquad \text{(Eq. 3.3.3)}$$

where T is defined in (Eq. 2.4.16), V is defined in (Eq. 2.4.17), and ω is defined by (Eq. 2.4.18). From (Eq. 3.3.3) it is easy to see $\theta_{\text{mx}} = -\omega$ and

$$f_{H1,\theta_{\text{mx}}}(x^n, y^n) = \frac{1}{(2\pi)^n N_o^n} \exp\left(-\frac{1}{2N_0}[T^2 + nE]\right) \exp\left(\frac{\sqrt{E}}{N_0} V\right) \qquad \text{(Eq. 3.3.4)}$$

Our Neyman Pearson like rule decides in favor of H_1 if

$$\log \frac{f_{H1,\theta_{\text{mx}}}}{f_{HO}} = -\frac{nE}{2N_o} + \frac{\sqrt{E}}{N_0} V > \gamma \qquad \text{(Eq. 3.3.5)}$$

Where γ is selected to satisfy the false alarm constraint. After simplification, our Neyman Pearson like rule reduces down to the exact same rule found in section 2.4. In other words, the optimal Neyman Pearson test is equivalent to a test which finds the maximum likelihood estimate of θ_o and then does the Neyman Pearson detection. Based on this observation we will design our robust detection rule as follows, first use robust estimation to estimate θ_o and then use an appropriate robust decision rule from Section 3.1.

3.3.1 Robust Estimation of the unknown initial phase

The maximum likelihood estimate of θ_o can be found by forming complex numbers from the data pairs $z_i = x_i + jy_i$ and forming complex numbers $h_i = \cos(p(i)) + j\sin(p(i))$, computing the complex metric $T = \sum_{i=1}^{n} z_i h_i^*$ and $\hat{\theta}$ is the angle of T, $\hat{\theta} = atan2(Im\{T\}/Re\{T\})$. This is not an M-estimate; however, the real and imaginary components of T are M-estimates. We therefore examine the components used in the metric. Let

$$R_i = -X_i \sin(p(i)) + Y_i \cos(p(i)) \qquad \text{(Eq. 3.3.6)}$$

$$S_i = X_i \cos(p(i)) + Y_i \sin(p(i)) \qquad \text{(Eq. 3.3.7)}$$

where all random variables are conditioned on θ. Then clearly

$$T = \sum\nolimits_{i=1}^{n} jR_i + \sum\nolimits_{i=1}^{n} S_i \qquad \text{(Eq. 3.3.8)}$$

Since R_i and S_i are linearly combinations of Gaussian variables, they also are Gaussian. In fact $\{R_i\}$ are Independent Identically Distributed Gaussian variables with mean $\sqrt{E}\sin(\theta_o)$ and variance N_o, and $\{S_i\}$ are Independent Identically Distributed Gaussian with mean $\sqrt{E}\cos(\theta_o)$ and variance N_o. We can now develop a robust estimate of θ_o, denoted $\hat{\theta}^*$, by independently using robust estimates of the means of $\{R_i\}$ and $\{S_i\}$.

The formal mathematical approach to finding the robust estimate of a location parameter (the mean) of a generating process within a specific class of functions $\mathfrak{J}$ is due to Huber and can be found in detail in *Detection and Estimation* (p320-30) [3]. This work will summarize the results and assumptions as they pertain towards robust estimate of the means of Gaussian variables.

Let us consider the possible occurrence of extreme independent outliers on independent identically distributed Gaussian data. Let $f_1(u)$ denote the density function of a zero-mean Gaussian variable with variance N_o. Let outliers appear with frequency ε and let their distribution be represented by some unknown density function g whose entropy and Fisher information measure exist, and which is such that $g(u) = g(-u)\ \forall\, u$; $\int_R g(u)[\log(g(u) + H(g)]^2\, du < \infty$; $\dot{g}(u) < \infty\ \forall\, u$, where $H(g)$ is the entropy of g and $\dot{g}(u) = \frac{\partial}{\partial u} g(u)$. Let $\mathfrak{J}$ be the class of all valid density functions which can be written

$$f(u) = (1-\varepsilon)f_1(u) + \varepsilon g(u) \qquad \text{(Eq. 3.3.9)}$$

where g can be any density function which satisfies the constraints given above. Given f in $\mathfrak{I}$, let $I(f)$ denote the Fisher information measure. Define f^* which if exists satisfies

$$I(f^*) = \underset{f \in \mathfrak{I}}{inf} I(f) \qquad \text{(Eq. 3.3.10)}$$

For $0 < \varepsilon < 0.03$, f^* is as follows

$$f^*(x) = \begin{cases} \frac{1}{\sqrt{N_o}} \phi\left(\frac{a}{\sqrt{N_o}}\right) \left[\cos\left(\frac{ca}{2N_o}\right)\right]^{-2} \left[\cos\left(\frac{cx}{2N_o}\right)\right]^{2} & -a \le x \le a \\ \frac{1}{\sqrt{N_o}} \phi\left(\frac{x}{\sqrt{N_o}}\right) & a < |x| < b \\ \frac{1}{\sqrt{N_o}} \phi\left(\frac{b}{\sqrt{N_o}}\right) \exp\left(-\frac{b}{N_o}|x-b|\right) & |x| \ge b \end{cases} \qquad \text{(Eq. 3.3.11)}$$

Where the constants a, b, c are such that

$$\begin{gathered} c \tan\left(\frac{ca}{2N_o}\right) = a \\ \phi\left(\frac{b}{\sqrt{N_o}}\right) = \frac{b}{\sqrt{N_o}}\left[\Phi\left(-\frac{b}{\sqrt{N_o}}\right) + \varepsilon\right] \\ \phi\left(\frac{a}{\sqrt{N_o}}\right)\left[\cos\left(\frac{ca}{2N_o}\right)\right]^{-2} = \frac{2c}{\sqrt{N_o}}\left[\frac{ca}{N_o} + \sin\left(\frac{ca}{N_o}\right)\right]^{-1}\left[\Phi\left(-\frac{a}{\sqrt{N_o}}\right) - 2^{-1} - \varepsilon\right] \end{gathered} \qquad \text{(Eq. 3.3.12)}$$

Let $\psi^* = -\dot{f}^*/f^*$. For $x \ge 0$,

$$\psi^*(x) = \begin{cases} \frac{c}{N_o}\tan\left(\frac{cx}{2N_o}\right) & 0 \le x \le a \\ \frac{x}{N_o} & a < x < b \\ \frac{b}{N_o} & x \ge b \end{cases} \qquad \text{(Eq. 3.3.13)}$$

and $\psi^*(x) = -\psi^*(-x)$. Given some data sequence x^n generated from some $f(u-m)$, where $f(u)$ is within $\mathfrak{I}$ and m is a constant, the robust estimation of m, denoted $\widehat{m}$ is

$$\widehat{m}^*(x^n): \sum_{i=1}^{n} \psi^*[x_i - \widehat{m}^*(x^n)] = 0 \qquad \text{(Eq. 3.3.14)}$$

This robust estimation $\widehat{m}^*(x^n)$ is consistent at every $f \in \mathfrak{J}$. Let $\beta(\psi, f)$ be the rate of convergence of some M-estimate ψ at f. Let $\wp$ be the class of all ψ such that $\psi(x) = -\psi(-x)$. The rate of convergence of our robust M-estimate is such that

$$\beta(\psi^*, f) \leq \beta(\psi^*, f^*) \leq \beta(\psi, f^*) \;\; \forall f \in \mathfrak{J};\ \forall \psi \in \wp \qquad \text{(Eq. 3.3.15)}$$

We define our robust estimate $\hat{\theta}$ as follows: use the estimation rule in (Eq. 3.3.13) to obtain estimates $\widehat{R^*}$ and $\widehat{S^*}$ and then define $\hat{\theta} = arctan2\left(\widehat{R^*}/\widehat{S^*}\right)$. At a given $f \in \mathfrak{J}$ as $n \to \infty$ the random variables $\widehat{\boldsymbol{R}^*}$ and $\widehat{\boldsymbol{S}^*}$ are asymptotically Gaussian.

We will consider

$$f(x_i, y_i) = (1-\varepsilon) f_1(x_i, y_i) + \varepsilon g_{x,y}(x_i, y_i) \qquad \text{(Eq. 3.3.16)}$$

where $g(x_i, y_i) = g_x(x_i) g_y(y_i)$ and where $g_x(x_i)$and $g_y(y_i)$ are defined via a discrete random variable q as follows

$$\begin{matrix} Prob\{q = 0\} = 0.5 \\ Prob\{q = 1\} = 0.5 \end{matrix} \qquad \text{(Eq. 3.3.17)}$$

$$g_x(x_i) = \begin{cases} \dfrac{1}{\sqrt{N_o}} \phi\left(\dfrac{x_i - \left[\sqrt{E}\cos(p(i) + \theta_o) + m\cos\left(p(i)\right)\right]}{\sqrt{N_o}} \right); q = 0 \\ \dfrac{1}{\sqrt{N_o}} \phi\left(\dfrac{x_i - \left[\sqrt{E}\cos(p(i) + \theta_o) - m\cos\left(p(i)\right)\right]}{\sqrt{N_o}} \right); q = 1 \end{cases} \qquad \text{(Eq. 3.3.18)}$$

$$g_y(y_i) = \begin{cases} \frac{1}{\sqrt{N_o}} \phi\left(\frac{y_i - \left[\sqrt{E}\sin(p(i)+\theta_o) + m\sin\big(p(i)\big)\right]}{\sqrt{N_o}}\right); q = 0 \\ \frac{1}{\sqrt{N_o}} \phi\left(\frac{y_i - \left[\sqrt{E}\sin(p(i)+\theta_o) - m\sin\big(p(i)\big)\right]}{\sqrt{N_o}}\right); q = 1 \end{cases} \qquad \text{(Eq. 3.3.19)}$$

Define $f_R(u) = (1-\varepsilon)f_{1,R}(u) + \varepsilon g_R(u)$ as the density function of the variable $\boldsymbol{R}$, where R is as defined in (Eq. 3.3.6) and $\{x_i, y_i\}$ are realizations of random variables defined by (Eq. 3.3.16). Also define $f_S(u) = (1-\varepsilon)f_{1,S}(u) + \varepsilon g_S(u)$ similarly where S is as defined in (Eq. 3.3.7), we obtain

$$g_R(u) = \frac{.5}{\sqrt{N_o}}\phi\left(\frac{u - (\sqrt{E}\sin(\theta_o) + m)}{\sqrt{N_o}}\right) + \frac{.5}{\sqrt{N_o}}\phi\left(\frac{u - (\sqrt{E}\sin(\theta_o) - m)}{\sqrt{N_o}}\right) \qquad \text{(Eq. 3.3.20)}$$

$$g_S(u) = \frac{0.5}{\sqrt{N_o}}\phi\left(\frac{u - (\sqrt{E}\cos(\theta_o) + m)}{\sqrt{N_o}}\right) + \frac{.5}{\sqrt{N_o}}\phi\left(\frac{u - (\sqrt{E}\cos(\theta_o) - m)}{\sqrt{N_o}}\right) \qquad \text{(Eq. 3.3.21)}$$

Clearly, $g_R\big(u - \sqrt{E}\sin(\theta_o)\big)$ and $g_s\big(u - \sqrt{E}\cos(\theta_o)\big)$ are symmetric density functions around zero; the outlier model is such that $f_R, f_s \in \mathfrak{I}$ where $\mathfrak{I}$ is defined in (Eq. 3.3.9). The robust estimates $\widehat{R^*}$ and $\widehat{S^*}$ at f_R and f_s are asymptotically Gaussian. Denote μ_R and σ_R as the mean and standard deviation of $\widehat{R^*}$ at f_R as $n \to \infty$. Define μ_S and σ_S similarly. By applying the theorem in Detection and Estimation (p 312) [3], we find

$$\mu_R = \sqrt{E}\sin(\theta_o) \qquad \text{(Eq. 3.3.22)}$$

$$\mu_S = \sqrt{E}\cos(\theta_o) \qquad \text{(Eq. 3.3.23)}$$

$$\sigma_R^2 = n^{-1} \frac{\int_{-\infty}^{\infty} [\psi^*(u-\mu_R)]^2 f_R(u) du}{\left[\frac{\partial}{\partial \tau} \int_{-\infty}^{\infty} \psi^*(u-\tau) f_R(u) du \Big|_{\tau=\mu_R} \right]^2} \qquad \text{(Eq. 3.3.24)}$$

$$\sigma_S^2 = n^{-1} \frac{\int_{-\infty}^{\infty} [\psi^*(u-\mu_S)]^2 f_S(u) du}{\left[\frac{\partial}{\partial \tau} \int_{-\infty}^{\infty} \psi^*(u-\tau) f_S(u) du \Big|_{\tau=\mu_S} \right]^2} \qquad \text{(Eq. 3.3.25)}$$

To simplify $\int_{-\infty}^{\infty} [\psi^*(u-\mu_R)]^2 f_R(u) du$ and $\int_{-\infty}^{\infty} [\psi^*(u-\mu_S)]^2 f_S(u) du$, we will define a new function

$$\omega_1(a, c, N_0, m) = \int_{-\frac{m}{\sqrt{N_o}}}^{\frac{a-m}{\sqrt{N_o}}} \frac{c^2}{N_o^2} \tan^2\left(\frac{c\left(y\sqrt{N_o} + m\right)}{2N_o} \right) \phi(y) dy \qquad \text{(Eq. 3.3.26)}$$

and we will also use the following equalities

$$\int_b^{\infty} \frac{b^2}{N_o^2} \frac{1}{\sqrt{N_o}} \phi\left(\frac{y-m}{\sqrt{N_o}} \right) dy = \frac{b^2}{N_o^2} \Phi\left(\frac{b-m}{\sqrt{N_o}} \right) \qquad \text{(Eq. 3.3.27)}$$

$$\begin{aligned} &\int_a^b \frac{y^2}{N_o^2} \frac{1}{\sqrt{N_o}} \phi\left(\frac{y-m}{\sqrt{N_o}} \right) dy \\ &\quad = \frac{m^2 + N_o}{N_o^2} \left[\Phi\left(\frac{b-m}{\sqrt{N_o}} \right) - \Phi\left(\frac{a-m}{\sqrt{N_o}} \right) \right] \qquad \text{(Eq. 3.3.28)} \\ &\quad + \frac{\sqrt{N_o}}{N_o^2} (a+m) \phi\left(\frac{a-m}{\sqrt{N_o}} \right) - \frac{\sqrt{N_o}}{N_o^2} (b+m) \phi\left(\frac{b-m}{\sqrt{N_o}} \right) \end{aligned}$$

We obtain

$\int_{-\infty}^{\infty}[\psi^*(u-\mu_S)]^2 f_S(u)du = \int_{-\infty}^{\infty}[\psi^*(u-\mu_R)]^2 f_R(u)du = 2(1-$

$\varepsilon\omega1a,c,N0,0+\varepsilon\omega1a,c,N0,m+21-\varepsilon NoNo2\Phi bNo-\Phi aNo+NoNo2a\phi aNo$

$-NoNo2b\phi bNo+\varepsilon m2+NoNo2\Phi b-mNo-\Phi a-mNo+NoNo2a+m\phi a-$

$mNo-NoNo2b+m\phi b-mNo+21-\varepsilon b2No2\Phi-bNo+\varepsilon b2No2\Phi-b+mNo$ (Eq. 3.3.29)

To simplify $\frac{\partial}{\partial\tau}\int_{-\infty}^{\infty}\psi^*(u-\tau)f_R(u)du\Big|_{\tau=\mu_R}$ and $\frac{\partial}{\partial\tau}\int_{-\infty}^{\infty}\psi^*(u-\tau)f_S(u)du\Big|_{\tau=\mu_s}$, we will define a new function

$$\omega_2(a,c,N_0,m) = \int_{-\frac{m}{\sqrt{N_o}}}^{\frac{a-m}{\sqrt{N_o}}} \frac{c^2}{2N_o^2}\sec^2\left(\frac{c\left(y\sqrt{N_o}+m\right)}{2N_o}\right)\phi(y)dy \qquad \text{(Eq. 3.3.30)}$$

We obtain

$$\begin{aligned}\frac{\partial}{\partial\tau}\int_{-\infty}^{\infty}\psi^*(u-\tau)f_R(u)du\Big|_{\tau=\mu_R} &= \frac{\partial}{\partial\tau}\int_{-\infty}^{\infty}\psi^*(u-\tau)f_S(u)du\Big|_{\tau=\mu_S} \\ &= 2(1-\varepsilon)\omega_2(a,c,N_0,0)+\varepsilon\omega_2(a,c,N_0,m) \\ &\quad +\frac{2(1-\varepsilon)}{N_o}\left[\Phi\left(\frac{b}{\sqrt{N_o}}\right)-\Phi\left(\frac{a}{\sqrt{N_o}}\right)\right] \qquad \text{(Eq. 3.3.31)} \\ &\quad +\frac{\varepsilon}{N_o}\left[\Phi\left(\frac{b-m}{\sqrt{N_o}}\right)-\Phi\left(\frac{a-m}{\sqrt{N_o}}\right)\right]\end{aligned}$$

We will now find the probability density function of our estimate $\widehat{\boldsymbol{\theta}^*}$ as $n\to\infty$. To simplify future expressions we will find the probability density function of $\boldsymbol{\vartheta}$, $f(\vartheta)$, where $\vartheta=\widehat{\theta^*}-\theta_o$. A realization of ϑ is found as the angle of $(\widehat{S^*}+j\widehat{R^*})\exp(-\mathrm{j}\theta_o)$. The complex random variable

$(\widehat{\boldsymbol{S}^*} + j\widehat{\boldsymbol{R}^*})\exp\,(-\mathrm{j}\theta_o)$ is Gaussian with both real and imaginary parts having variance $\sigma_S^2 = \sigma_R^2$ and whose real mean is $\sqrt{\mu_S^2 + +\mu_R^2} = \sqrt{E}$ and whose imaginary mean is zero. Define

$$z_n = Im\{(\widehat{S^*} + j\widehat{R^*})\exp(-\mathrm{j}\theta_o)\} \qquad \text{(Eq. 3.3.32)}$$

$$z_d = Re\{(\widehat{S^*} + j\widehat{R^*})\exp(-\mathrm{j}\theta_o)\} \qquad \text{(Eq. 3.3.33)}$$

where $Im\{x\}$ and $Re\{x\}$ denote taking the imaginary and real part of x respectively. Define

$$Z = \frac{Im\{(\widehat{S^*} + j\widehat{R^*})\exp(-\mathrm{j}\theta_o)\}}{Re\{(\widehat{S^*} + j\widehat{R^*})\exp(-\mathrm{j}\theta_o)\}} = \frac{z_n}{z_d} \qquad \text{(Eq. 3.3.34)}$$

We find

$$\vartheta = arctan2(Z) = \begin{cases} \arctan(Z)\,; & if\ z_d > 0 \\ \arctan(Z) + \pi; & if\ z_n < 0 \end{cases} \qquad \text{(Eq. 3.3.35)}$$

Let $f_{z|z_d^+}$ and $f_{z|z_d^-}$ denote the probability density function of $\boldsymbol{Z}$ conditioned on $z_d > 0$ and $z_d < 0$ respectively. Let f_ϑ denote the probability density function of the random variable $\boldsymbol{\vartheta}$. We obtain

$$f_\vartheta(\vartheta) = \begin{cases} (1 + \tan^2\vartheta) f_{z|z_d^+}(\tan(\vartheta)) Prob\{z_d > 0\}; & -\dfrac{\pi}{2} < \vartheta < \dfrac{\pi}{2} \\ (1 + \tan^2\vartheta) f_{z|z_d^-}(\tan(\vartheta)) Prob\{(z_d < 0)\}; & \dfrac{\pi}{2} < \vartheta < \dfrac{3\pi}{2} \end{cases} \qquad \text{(Eq. 3.3.36)}$$

$$f_{z|z_d^+}(\tan(\vartheta))Prob\{z_d > 0\}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\cos(\vartheta)\exp\left(-\frac{Esin^2(\vartheta)}{2\sigma^2}\right)\left[\sigma\cos(\vartheta)\,\phi\left(\frac{\sqrt{E}\cos(\vartheta)}{\sigma}\right)\right.$$
$$\left. + \sqrt{E}\cos^2(\vartheta)\Phi\left(\frac{\sqrt{E}\cos(\vartheta)}{\sigma}\right)\right] \tag{Eq. 3.3.37}$$

$$f_{z|z_d^-}(\tan(\vartheta))Prob\{z_d > 0\}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}|\cos(\vartheta)|\ \exp\left(-\frac{Esin^2(\vartheta)}{2\sigma^2}\right)\left[\sigma|\cos(\vartheta)|\ \phi\left(\frac{\sqrt{E}\cos(\vartheta)}{\sigma}\right)\right.$$
$$\left. - \sqrt{E}\cos^2(\vartheta)\Phi\left(\frac{-\sqrt{E}\,|\cos(\vartheta)|}{\sigma}\right)\right] \tag{Eq. 3.3.38}$$

Using (Eq. 3.3.37) and (Eq. 3.3.38) we can write (Eq. 3.3.36) as

$$f_\vartheta(\vartheta) = \phi\left(\frac{\sqrt{E}\sin(\vartheta)}{\sigma}\right)\left[\sigma\phi\left(\frac{\sqrt{E}\cos(\vartheta)}{\sigma}\right) + \sqrt{E}\,cos(\vartheta)\Phi\left(\frac{\sqrt{E}\cos(\vartheta)}{\sigma}\right)\right] \tag{Eq. 3.3.39}$$

The probability density function of $\widehat{\theta^*}$ as $n \to \infty$ is

$$f_{\widehat{\theta^*}}(x) = f_\vartheta(x - \theta_o) \tag{Eq. 3.3.40}$$

3.3.2 Robust Detection

As noted earlier, when the initial phase is unknown (Section 2.4), the optimal detection at the nominal model is equivalent to finding the maximum likelihood estimate for θ_o. We denote this estimate $\hat{\theta}$, assuming $\theta_o = \hat{\theta}$ and examining the same metric used in optimal detection at the nominal model where the waveform is completely known (Section 2.2) . In similar fashion, based on our outlier model, we can use the appropriate robust detection metric found in sections 3.1.1 or 3.1.2 under the assumption that $\theta_o = \widehat{\theta^*}$, where $\widehat{\theta^*}$ is obtained from the method described in 3.3.1. We will assume an outlier model where outliers occur per data pairs (x_i, y_i) and therefore, we will adopt the detection rule in section 3.1.1. We will focus our analysis on the case where the occurrence of outliers is the same under both H_o and H_1 and is less than 0.5 ($\varepsilon = \varepsilon_0 = \varepsilon_1 < 0.5$). Define

$$k_i(\widehat{\theta^*}) = x_i cos\left(p_i + \widehat{\theta^*}\right) + y_i sin\left(p_i + \widehat{\theta^*}\right) \qquad \text{(Eq. 3.3.41)}$$

$$z(k_i) = \begin{cases} \mathrm{d}_1 & k_i \leq d \\ k_i & d_1 < k_i < d_0 \\ \mathrm{d}_0 & k_i \geq \sqrt{E} - d \end{cases} \qquad \text{(Eq. 3.3.42)}$$

where d is the solutions to (Eq. 3.1.19). The robust decision rule is

$$\delta^*(k^n) = \begin{cases} 1 \quad k^n \colon n^{-1} \sum_{i=1}^{n} z(k_i) > \lambda^* \\ r^* \quad k^n \colon n^{-1} \sum_{i=1}^{n} z(k_i) = \lambda^* \\ 0 \quad k^n \colon n^{-1} \sum_{i=1}^{n} z(k_i) < \lambda^* \end{cases} \qquad \text{(Eq. 3.3.43)}$$

where $z(k_i)$ is as in (Eq. 3.3.42) and where λ^* and r^* are such that

$$\Pr\left\{ n^{-1} \sum_{i=1}^{n} z(k_i) > \lambda^* \middle| f_0^* \right\} + r^* \Pr\left\{ n^{-1} \sum_{i=1}^{n} z(k_i) = \lambda^* \middle| f_0^* \right\} = \alpha \qquad \text{(Eq. 3.3.44)}$$

The least favorable density functions f_0^* and f_1^* are given by (Eq. 3.1.16) and (Eq. 3.1.17). The density function of $n^{-1} \sum_{i=1}^{n} Z(k_i)$, conditioned on f_0^* and some $\widehat{\theta^*}$, is the same regardless of the value of $\widehat{\theta^*}$. Therefore the density function of $n^{-1} \sum_{i=1}^{n} Z(k_i)$ conditioned on f_0^* is the same as the density function of $n^{-1} \sum_{i=1}^{n} Z(k_i)$ conditioned on f_0^* and $\widehat{\theta^*} = \theta_o$. From section 3.1.1, this density function is asymptotically Gaussian with mean and variance given by (Eq. 3.1.23) and (Eq. 3.1.24), respectively. Denote by $m_{f_0^*}$ and $v_{f_0^*}$ respectively the mean and variance of $n^{-1}\sum_{i=1}^{n} Z(k_i)$ conditioned on f_0^*.

For asymptotically large values of $n,$ the constant r^* in (Eq. 3.3.43) can be taken equal to zero; we thus then obtain

$$\lambda^* = m_{f_0^*} + v_{f_0^*}\, \Phi^{-1}(1-\alpha) \qquad \text{(Eq. 3.3.45)}$$

Denote by $m_{f_1}(\widehat{\theta^*})$ and $v_{f_1}^2(\widehat{\theta^*})$ respectively the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$, conditioned on f_1 and $\widehat{\theta^*}$, for large n values. Since $\{K_i\}$ conditioned on f_1 and $\widehat{\theta^*}$ are independent identically distributed Gaussian variables with mean $\sqrt{E}\cos\left(\widehat{\theta^*}-\theta_o\right)$ and variance N_o, $m_{f_1}(\widehat{\theta^*})$ and $v_{f_1}^2\left(\widehat{\theta^*}\right)$ are respectively equal to $m_{g_m}(\mu)$ and $v_{g_m}^2(\mu)$ in (Eq. 3.1.29) and (Eq. 3.1.30) with $\mu = \frac{\sqrt{E}}{\sqrt{\mathrm{N_o}}}\cos\left(\widehat{\theta^*}-\theta_o\right)$. Let

$$f = (1-\varepsilon) f_1 + \varepsilon g_m \qquad \text{(Eq. 3.3.46)}$$

where

$$\begin{aligned} g_m(k) = {} & \frac{.5}{\sqrt{N_o}}\phi\left(\frac{k-(\sqrt{E}\cos\left(\widehat{\theta^*}-\theta_o\right)+m\cos\left(\widehat{\theta^*}\right))}{\sqrt{N_o}}\right) \\ & + \frac{.5}{\sqrt{N_o}}\phi\left(\frac{k-(\sqrt{E}\cos\left(\widehat{\theta^*}-\theta_o\right)-m\cos\left(\widehat{\theta^*}\right))}{\sqrt{N_o}}\right) \end{aligned} \qquad \text{(Eq. 3.3.47)}$$

Equivalently, $g_m(k)$ is the density resulting from the density functions $g_{x_i}(x_i)$ and $g_{y_i}(y_i)$ in (Eq. 3.3.18) and (Eq. 3.3.19). Define

$$g_{ma}(k) = \frac{1}{\sqrt{N_o}}\phi\left(\frac{k-(\sqrt{E}\cos\left(\widehat{\theta^*}-\theta_o\right)+m\cos\left(\widehat{\theta^*}\right))}{\sqrt{N_o}}\right) \qquad \text{(Eq. 3.3.48)}$$

$$g_{mb}(k) = \frac{1}{\sqrt{N_o}}\phi\left(\frac{k - (\sqrt{E}\cos(\widehat{\theta^*} - \theta_o) - mcos(\widehat{\theta^*}))}{\sqrt{N_o}}\right) \qquad \text{(Eq. 3.3.49)}$$

We rewrite (Eq. 3.3.46) as,

$$f = (1-\varepsilon)f_1 + \frac{\varepsilon}{2}g_{ma} + \frac{\varepsilon}{2}g_{mb} \qquad \text{(Eq. 3.3.50)}$$

Denote by $m_{g_{ma}}(\widehat{\theta^*})$ and $v_{g_{ma}}^2(\widehat{\theta^*})$ respectively the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$, conditioned on g_{ma} and $\widehat{\theta^*}$, for large values of n . Define $m_{g_{mb}}(\widehat{\theta^*})$ and $v_{g_{mb}}^2(\widehat{\theta^*})$ similarly. Since $g_{ma}(k)$ is a Gaussian probability density function with variance N_0, $m_{g_{ma}}(\widehat{\theta^*})$ and $v_{g_{ma}}^2(\widehat{\theta^*})$ are respectively equal to $m_{g_m}(\mu)$ and $v_{g_m}^2(\mu)$ in (Eq. 3.1.29) and (Eq. 3.1.30), with $\mu = \frac{\sqrt{E}}{\sqrt{\mathrm{N_o}}}\cos(\widehat{\theta^*} - \theta_o) + mcos(\widehat{\theta^*})$. Similarly, since $g_{mb}(k)$ is a Gaussian probability density function with variance N_0, $m_{g_{mb}}(\widehat{\theta^*})$ and $v_{g_{mb}}^2(\widehat{\theta^*})$ are respectively equal to $m_{g_m}(\mu)$ and $v_{g_m}^2(\mu)$ in (Eq. 3.1.29) and (Eq. 3.1.30), with $\mu = \frac{\sqrt{E}}{\sqrt{\mathrm{N_o}}}\cos(\widehat{\theta^*} - \theta_o) - mcos(\widehat{\theta^*})$.

Denote by $m_f(\widehat{\theta^*})$ and $v_f^2(\widehat{\theta^*})$ respectively the mean and variance of the random variable $n^{-1}\sum_{i=1}^{n} z(K_i)$, conditioned on f in (Eq. 3.3.50) and $\widehat{\theta^*}$, for large values of n . We obtain

$$m_f(\widehat{\theta^*}) = (1-\varepsilon)m_{f_1}(\widehat{\theta^*}) + \frac{\varepsilon}{2}m_{g_{ma}}(\widehat{\theta^*}) + \frac{\varepsilon}{2}m_{g_{mb}}(\widehat{\theta^*}) \qquad \text{(Eq. 3.3.51)}$$

$$\begin{aligned} v_f^2\left(\widehat{\theta^*}\right) = {} & (1-\varepsilon)v_{f_1}^2\left(\widehat{\theta^*}\right) + \frac{\varepsilon}{2} v_{g_{mb}}^2\left(\widehat{\theta^*}\right) + \frac{\varepsilon}{2} v_{g_{mb}}^2\left(\widehat{\theta^*}\right) \\ & + n^{-1}\left[(1-\varepsilon)\frac{\varepsilon}{2}\left(m_{f_1}\left(\widehat{\theta^*}\right) - m_{g_{ma}}\left(\widehat{\theta^*}\right)\right)^2 \right. \\ & + (1-\varepsilon)\frac{\varepsilon}{2}\left(m_{f_1}\left(\widehat{\theta^*}\right) - m_{g_{mb}}\left(\widehat{\theta^*}\right)\right)^2 \\ & \left. + \frac{\varepsilon^2}{4}\left(m_{g_{mb}}\left(\widehat{\theta^*}\right) - m_{g_{mb}}\left(\widehat{\theta^*}\right)\right)^2\right] \end{aligned} \qquad \text{(Eq. 3.3.52)}$$

Let us denote by $P_{n,\widehat{\theta^*}}(f)$ the power induced by the robust decision rule in (Eq. 3.3.43) at the density function f conditioned on $\widehat{\theta^*}$.

Due to (Eq. 3.3.45), we conclude: for $n \to \infty$,

$$P_{n,\widehat{\theta^*}}(f) = \Phi\left(\frac{m_f\left(\widehat{\theta^*}\right) - m_{f_0^*} - v_{f_o^*}\,\Phi^{-1}(1-\alpha)}{v_f\left(\widehat{\theta^*}\right)}\right) \qquad \text{(Eq. 3.3.53)}$$

Let us denote by $P_n(f)$ the power induced by the robust decision rule in (Eq. 3.3.43) at the density function f in (Eq. 3.3.50). From our analysis it is difficult to find $P_n(f)$, since $\widehat{\theta^*}$ is directly computed from our observed data (x^n, y^n). If $\widehat{\theta^*}$ was completely independent of the data sequence (x^n, y^n) and still had the distribution given by (Eq. 3.3.40), the theorem of total probability would give:

$$\begin{aligned} & P_n\left(f \middle| \widehat{\theta^*}\ selected\ independent\ of\ x^n, y^n\right) \\ & \qquad = \int_{-\pi+\theta_0}^{\pi+\theta_0} P_{n,\widehat{\theta^*}}(f) f_{\widehat{\theta^*}}(x)dx \end{aligned} \qquad \text{(Eq. 3.3.54)}$$

where $f_{\widehat{\theta^*}}(x)$ is defined in (Eq. 3.3.40). As $n \to \infty$ any single data pair (x_i, y_i) contribution to $\widehat{\theta^*}$ is negligible; therefore the distribution of $z(K_i)$ conditioned on a $\widehat{\theta^*}$ selected using the data pair

(x_i, y_i), is close to the distribution of $z(K_i)$ conditioned on a $\widehat{\theta^*}$ selected independently of (x_i, y_i). This logic cannot be extended to $n^{-1}\sum_{i=1}^{n} z(K_i)$.

3.3.3 Comparing robust detection with non-robust detection

Without a straight forward way to compute $P_n(f)$, even asymptotically, we will compare the performance of the robust and the optimal detectors, both under the assumption that the angle estimates, $\widehat{\theta^*}$ and $\hat{\theta}$, were selected independently of the data sequence (x^n, y^n). One way to achieve this independence is to assume that the analyst receives two separate independent sequences of data pairs from the same generating process, each of length n. The analyst uses the first to estimate $\widehat{\theta^*}$ or $\hat{\theta}$, and then based on this estimate uses the second to decide H_1 or H_0. We conjecture that this comparison will be at least qualitatively similar to the comparison of the case where the angle estimates are made directly from the same data sequence used to decide H_1 or H_0.

The $\widehat{\theta^*}$ and $\hat{\theta}$ estimates can be expressed as

$$\hat{\theta} = arctan2\left(\frac{\frac{1}{n}\sum_{i=1}^{n} R_i}{\frac{1}{n}\sum_{i=1}^{n} S_i}\right) \qquad \text{(Eq. 3.3.55)}$$

$$\widehat{\theta^*} = arctan2\left(\frac{\frac{1}{n}\sum_{i=1}^{n} \widehat{R_i}}{\frac{1}{n}\sum_{i=1}^{n} \widehat{S_i}}\right) \qquad \text{(Eq. 3.3.56)}$$

Denote by $f_1^*(m)$ the distribution defined in (Eq. 3.3.16). We note that $f_1^*(0) = f_1$, the nominal distribution. Let us denote by $\theta(m)$ the probability distribution of $\hat{\theta}$. Define $\theta^*(m)$ similarly. Under f_1 and f_1^*, as $n \to \infty$, the variables $\frac{1}{n}\sum_{i=1}^{n} R_i$, $\frac{1}{n}\sum_{i=1}^{n} S_i$, $\frac{1}{n}\sum_{i=1}^{n} \widehat{R_i}$, and $\frac{1}{n}\sum_{i=1}^{n} \widehat{S_i}$ are Gaussian; therefore, the $\theta(m)$, and $\theta^*(m)$ are of the mathematical form found in (Eq. 3.3.40).

We find that $\theta(m)$, and $\theta^*(m)$ both have the same value of E but generally different σ. Let us denote by $\sigma(m)$ and $\sigma^*(m)$ the σ parameter of $\theta(m)$ and $\theta^*(m)$ respectively. Figure 3-16, shows $\sigma(m)$ and $\sigma^*(m)$ for various values of m where $N_o = 10,\ n = 100,$ and $\varepsilon = 0.001.$ When $m = 0$, $\sigma(m) < \sigma^*(m)$ shows that the optimal estimate at f_1 is superior. However, as m increases, $\sigma(m)$ increases without bound, whereas $\sigma^*(m)$ does not; $\theta^*(m)$ does not change significantly as a function of m. Figure 3-17 shows $\sigma(0)$ and $\sigma^*(0)$ as a function of ε at f_1 where $N_0 = 10$, and $n = 100$; as ε increases $\sigma^*(0)$ increases showing the tradeoff between protecting against a greater frequency of outliers and optimal performance at f_1. Figure 3-18, shows $\theta(0)$ and $\theta^*(0)$ where $E = 1, N_0 = 10, \varepsilon = 0.03$ and $n = 100$; the closeness of the two distributions indicates that the performance of robust estimate is close to the performance of the optimal estimate at f_1. Figure 3-19, shows $\theta(100)$ and $\theta^*(100)$ where $E = 1, N_0 = 10, \varepsilon = 0.03$ and $n = 100$; the difference in the two distributions shows the danger of using the optimal estimate at f_1 with outliers present.

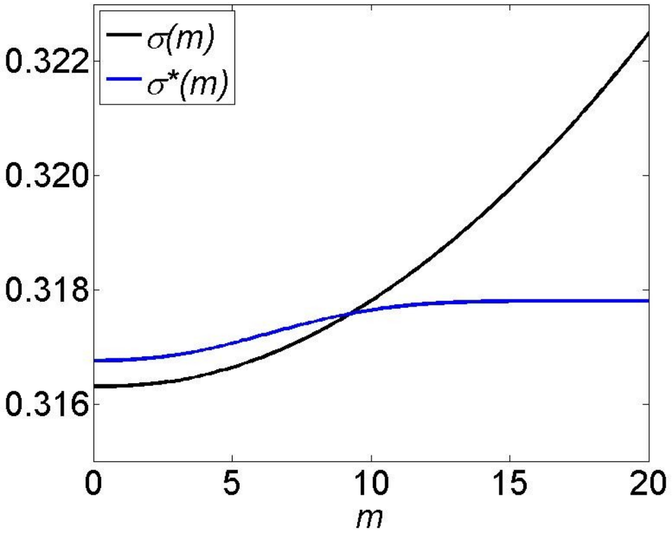

Figure 3-16: Sigma comparison where $N_o = 10,\ n = 100,$ and $\varepsilon = 0.001.$

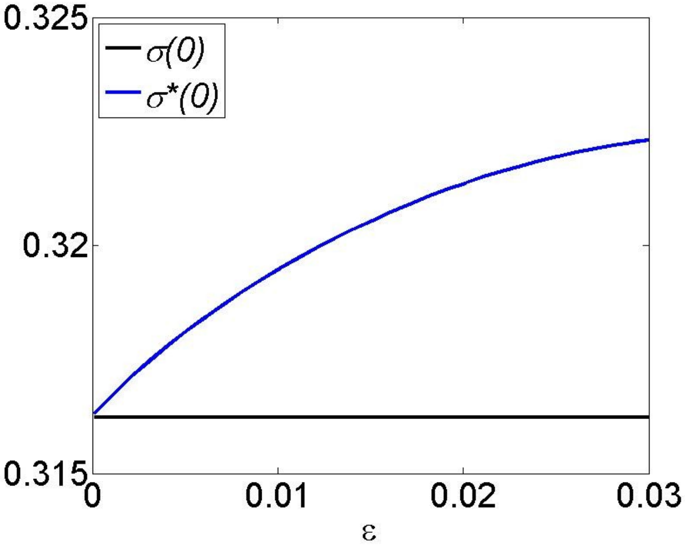

Figure 3-17: Sigma comparison at f_1 where $N_0 = 10$, and $n = 100$.

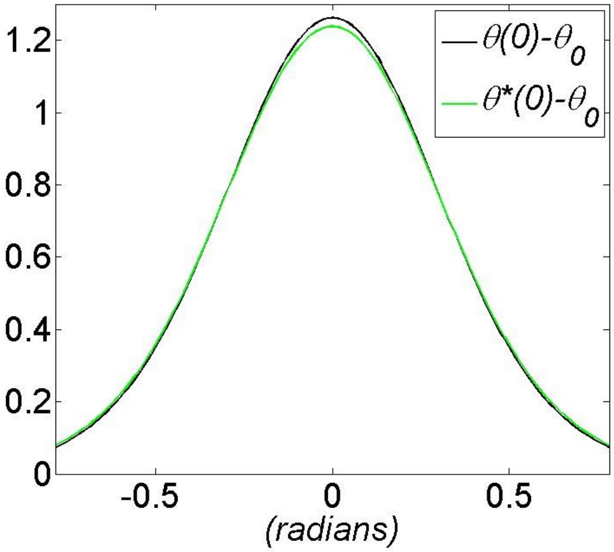

Figure 3-18: Probability density functions at f_1 where $E = 1, N_0 = 10, \varepsilon = 0.03$ and $n = 100$.

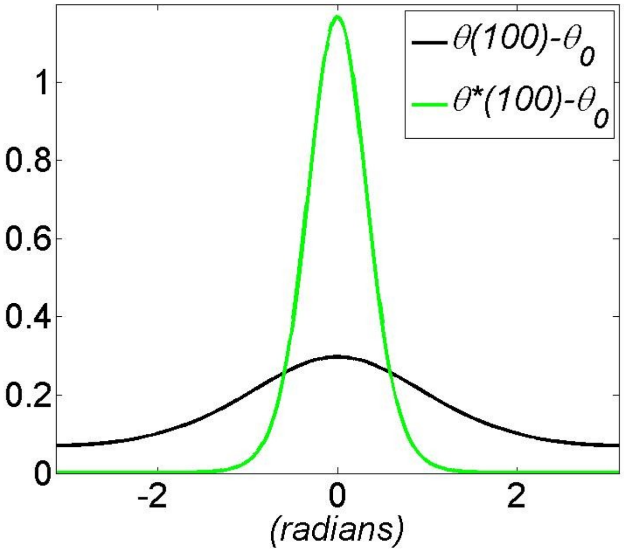

Figure 3-19: Probability density functions at f_1^* with $m = 100$, where $E = 1, N_0 = 10$, $\varepsilon = 0.03$ and $n = 100$.

Denote by $P^*(f_0^*, m)$ the power of the robust decision rule obtained by independently estimating $\widehat{\theta^*}$ and then using the robust decision rule given by (Eq. 3.3.43) at the density pair $[f_0^*, f_1^*(m)]$; $P^*(f_0^*, m)$ is found using the integral found in (Eq. 3.3.54). Denote by $P(f_o, m)$ the power of the robust decision rule obtained by independently estimating $\widehat{\theta}$ and then using the robust decision rule given by (Eq. 2.2.1) at the density pair $[f_o, f_1^*(m)]$; $P(f_o, m)$ is found using an integral with the same form as (Eq. 3.3.54), where the robust conditional power and robust probability density replace their the non-robust equivalents. Figure 3-20 shows $P^*(f_0^*, m)$ and $P(f_o, m)$ where $m = 0, E = 1,\ N_0 = 10,\ \varepsilon = 0.03$ and $\alpha = 1e - 5$ as a function of n. The difference between the two curves shows the performance loss induced by the robust decision rule designed at the highest permissible frequency of outliers ($\varepsilon = 0.03$) as compared to the optimal decision rule, both in the absence of outliers. Figure 3-21 shows $P^*(f_0^*, m)$ and $P(f_o, m)$ as functions of , where $m = 100, E = 1,\ N_0 = 10,\ \varepsilon = 0.03$ and $\alpha = 1e - 5$. The difference between the two curves shows the performance loss from using the optimal decision rule in the presence of outliers as compared to robust decision rule.

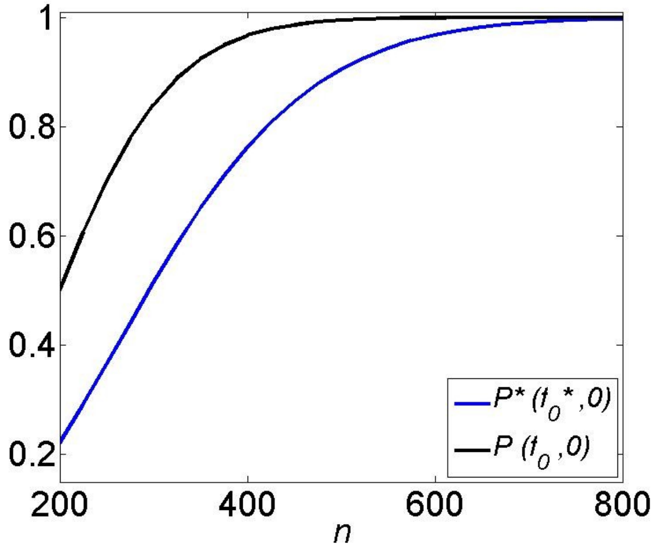

Figure 3-20: Power curves at $f_1^*(0) = f_1$ where $E = 1,\ N_0 = 10,\ \varepsilon = 0.03$ and $\alpha = 1e - 5$.

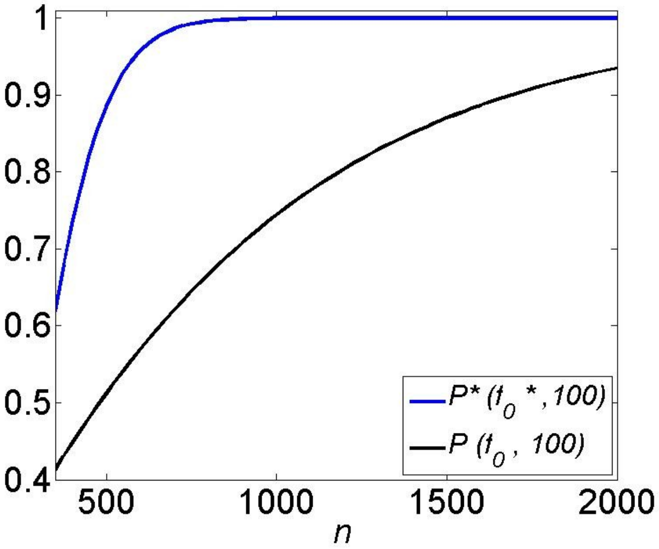

Figure 3-21: Power curves at $f_1^*(100)$ where $E = 1,\ N_0 = 10,\ \varepsilon = 0.03$ and $\alpha = 1e - 5$.

4 Thesis Conclusions

4.1 Summary of Findings

In Section 1, we developed the Neyman-Pearson detection rules for an FSK modulated signal that is transmitted via a channel modeled by additive white Gaussian noise and is filtered by an ideal brick-wall filter, for three cases: 1) when the waveform is known completely (coherent detection), 2) when the waveform is known except for a random initial phase (incoherent detection) and 3) when only the energy of the waveform is known (energy detection). To start, we used only first and second order statistics to compare the three detectors. Then, we compared the three detectors using full statistical descriptions of the involved waveforms. While the first and second order statistical analysis led to simple and often quoted performance comparison figures, its results were expectedly highly inferior to those induced by the full statistical analysis. The full statistical analysis showed that the performance of the three detectors is generally closer than the indications induced by the first and second order statistics analysis.

In Section 2, we developed the robust detectors for the three cases cited above, where the classes for both the null and emphatic hypothesis are the Huber classes. For the coherent detection case, we considered two outlier models: 1) when outliers occur over complex data pairs and 2) when outliers occur independently per datum. We derived analytic expressions for their asymptotic performance, when the outliers are drawn from a Gaussian distribution. Finally, for various models, we compared their performance to that of the optimal at the nominal distribution pair (the non-robust detector) detector, as the latter was developed in the previous section. For the energy detection case, we similarly derived the robust detector and compared it to the non-robust detector found in the previous section. In general, in the absence of outliers, our analysis showed slightly reduced performance for the robust detectors, as compared to that of the non-robust detectors. In contrast, in the presence of outliers, our analysis showed significantly reduced performance for the non-robust detectors, as compared to that of the robust detectors. Both the coherent and the energy detection generating processes considered were memoryless; as a result, the robust detection schemes were developed using existing theorems and are solutions to saddle-point optimization problems. However, the incoherent detection generating process is not memoryless and no known theorems could be used to derive the optimal detection scheme. Despite this, we were able to develop an intuitively pleasing incoherent robust detector for the incoherent case. Our robust detector first used the data to derive a robust initial phase estimate; using the latter phase estimate, our detector subsequently used the robust detection established methods for the coherent case. Due to the inter-dependence between the phase estimate and the subsequent detection, we did not find analytic expressions for the asymptotic performance of our detector. We derived analytic expressions for the asymptotic performance of a slightly modified incoherent robust detector, however, which removed the inter-dependence between the phase estimate and the detection steps. We compared the latter analytic expressions to non-robust versions, in both the presence and the absence of outliers. In the absence of outliers, these comparisons showed a slight performance loss when the robust detector is used, as compared to the case when the non-robust detector is used, instead. In the presence of outliers, the robust detector significantly outperformed the non-robust detector.

While our analysis focused specifically on the detection of FSK signals, Section 1 highlighted the importance of using full statistical models, when different detectors are compared and; Section 3 highlighted the danger of relaying on non-robust detectors in the presence of outliers. Non-robust detectors relay on complete knowledge of the underlying data generating distributions. In general, for all real-world problems, full knowledge of the data generating processes is nonfeasible. Our analysis may derive the generating distributions from approximate physics-based models. Otherwise, our analysis may attempt to estimate these distributions using data observations. In either case, the generating distributions are only approximately known. For the latter case, in particular, the tails of the distributions are difficult to estimate correctly. Robust detection performs well in the presence of outliers which occur at extreme values represented by the tails of our nominal distribution. In contrast, non-robust detection performance breaks down in the presence of even infrequent outliers. We therefore conclude that robust detectors should always be used over non-robust detectors, for all real-world problems.

4.2 Future Research

There are several areas to expand and further the research presented in this thesis. Future work could examine the non-robust and robust detection schemes of non FSK signals. While much of the results in this thesis extend to non-FSK band-limited signals with constant energy, many of the results need to be modified to address signals with time-varying energy such as Amplitude Modulated signals. In addition, this thesis examined detection under the assumption of an ideal brick-wall filter; future work could examine the detection schemes of signals using other filter models. Of particular interest are filters which are found in receiver hardware such as Finite Impulse Response filters. The current thesis focused on Neyman Pearson detection schemes where the number of samples for a given detection scheme is fixed. Future work could examine the performance of Sequential Detection Testing for both the non-robust and robust cases. The robust detection schemes performed data truncation based on the frequency of outliers. For most real-world applications, the frequency of outliers will not be known a priori, and may not be constant. Future work could examine methods to optimally estimate this frequency, and update this estimate as needed. Future work could also examine performance of the various detection schemes on simulated and real data. This performance could be compared to the analytic performance presented in this thesis. Some of the performance curves found in this thesis were based on the assumption of a large number of samples. Future work could do performance analysis assuming a small number of samples. While this thesis examined detection of FSK signals, future work could examine the optimal schemes to demodulate FSK signals with particular interest in robust demodulation algorithms. Finally, future research could examine ways to make the FSK signals more robust to extreme outliers, thereby increasing performance of robust detection demodulation algorithms.

BIBLIOGRAPHY

[1] P. Papantoni-Kazakos and I. M. Paz, "The Performance of a Digital FM System with Discriminator: Intersymbol Interference Effects," *IEEE Transactions on Communications,* vol. 23, no. 9, pp. 867-78, September 1975.

[2] J. G. Proakis and M. Salehi, Digital Communications, New York: McGraw-Hill, 2008, pp. 212-14.

[3] D. Kazakos and P. Papantoi-Kazakos, Detection and Estimation, New York: Computer Science Press, 1990.

[4] T. Papantoni, P. Papantoni-Kazakos and A. Burrell, "Robust Sequential Algorithms For the Detection of Changes In Data Generating Processes," University of Colorado Denver, Denver, 2009.

Appendix A : Extension of Huber Robust Detection

Let $\mathfrak{I}_1^n$ and $\mathfrak{I}_o^n$ be two disjoint classes of n-dimensional density functions, and let $\mathfrak{I}_1^n$ and $\mathfrak{I}_o^n$ represent respectively hypothesis H_1 and H_o. Denote by $\{f_i^n \colon f_{i,1}, \ldots, f_{i,n}\}$ a set of density functions within the class $\mathfrak{I}_i^n$. Assume that x^n is generated by a single set of density functions in either $\mathfrak{I}_1^n$ or $\mathfrak{I}_o^n$, denoted f_i^n. Finally, assume that $\mathfrak{I}_1^n$ and $\mathfrak{I}_o^n$ contain only memoryless processes. Then, given n, an observation vector x^n, and some f_i^n we have

$$f_i(x^n) = \prod_{j=1}^{n} f_{i,j}(x_i) \qquad \text{(Eq. A.1)}$$

Given some pair (f_o^n, f_1^n), our non-robust decision rule is

$$\delta(x^n) = \begin{cases} 1 & x^n \colon T(x^n) = n^{-1} \sum_{i=1}^{n} \log\left(\frac{f_{1,i}(x_i)}{f_{o,i}(x_i)}\right) > \lambda^* \\ r^* & x^n \colon T(x^n) = n^{-1} \sum_{i=1}^{n} \log\left(\frac{f_{1,i}(x_i)}{f_{o,i}(x_i)}\right) > \lambda^* \\ 0 & x^n \colon T(x^n) = n^{-1} \sum_{i=1}^{n} \log\left(\frac{f_{1,i}(x_i)}{f_{o,i}(x_i)}\right) > \lambda^* \end{cases} \qquad \text{(Eq. A.2)}$$

where r^*, λ^* are selected to satisfy the false alarm constraint. If some pair (f_o^{*n}, f_1^{*n}) exists, with decision rule $\delta^*(x^n)$ such that

$$P(\delta, f_1^{*n}) \leq P(\delta^*, f_1^{*n}) \leq P(\delta^*, f_1^n) \quad \forall f_1^n \in \mathfrak{I}_1^n \qquad \text{(Eq. A.3)}$$

and

$$P(\delta^*, f_0^n) \leq P(\delta^*, f_0^*) = \alpha \quad \forall f_0^n \in \mathfrak{I}_0^n \qquad \text{(Eq. A.4)}$$

then $\delta^*(x^n)$ is the robust decision rule.

Lemma A.1: Let there exist a least favorable pair $(f_{o,i}^*, f_{1,i}^*)$ for every $i: 1 \ldots n$, under a single sample hypothesis test, let $\Im_1^n$ and $\Im_o^n$ contain only memoryless processes and let each least favorable pair satisfy

$$\Pr\left\{log\left(\frac{f_{1,i}^*(X)}{f_{o,i}^*(X)}\right) < y \middle| f_{o,i}\right\} \geq \Pr\left\{log\left(\frac{f_{1,i}^*(X)}{f_{o,i}^*(X)}\right) < y \middle| f_{o,i}^*\right\}$$

$$\geq \Pr\left\{log\left(\frac{f_{1,i}^*(X)}{f_{o,i}^*(X)}\right) < y \middle| f_{1,i}^*\right\} \qquad \text{(Eq. A.5)}$$

$$\geq \Pr\left\{log\left(\frac{f_{1,i}^*(X)}{f_{o,i}^*(X)}\right) < y \middle| f_{1,i}\right\} \quad \forall\, i: 1, \ldots, n, \forall f_{1,i} \in \Im_1^n, \forall f_{0,i} \in \Im_0^n, \forall\, y \; in \; \Re$$

Then, the pair $(f_{o,i}^{*n}, f_{1,i}^{*n})$ is the least favorable density pair over $\Im_1^n$ and $\Im_o^n$. In other words, if a process is memoryless, the least favorable density pair can be found by taking the least favorable density pair over each data sample, conditioned on each least favorable density pair satisfying (Eq. A.5).

Proof: The condition in (Eq. A.5) implies that, given $f_{0,i} \in \Im_{0,i}$ and $f_{1,i} \in \Im_{1,i}$, a random variable Y and non-decreasing functions $g_{o,i}(y), g_{o,i}^*(y), g_{1,i}^*(y)$ and $g_{1,i}(y)$ exist, such that $g_{o,i}(y) \geq g_{o,i}^*(y) \geq g_{1,i}^*(y) \geq g_{1,i}(y) \;\; \forall\, y \text{ in } \Re$, and the functions $g_{o,i}(y), g_{o,i}^*(y), g_{1,i}^*(y),\; g_{1,i}(y)$ coincide respectively with the probabilities found in (Eq. A.5). The proof then follows the proof found in *Detection and Estimation* (Lemma 6.3.1, p.172-3) [3], replacing instances of $g_o(y), g_o^*(y), g_1^*(y),\; g_1\,(y)$ with $g_{o,i}(y), g_{o,i}^*(y), g_{1,i}^*(y),\; g_{1,i}(y)$.

If $\Im_{0,i}$ and $\Im_{1,i}$ are defined as Huber's classes, where Huber's classes are defined as

$$\Im_{o,i} = \{f_i(x) = (1-\varepsilon_o) f_{o,i}(x) + \varepsilon_o h(x) \;\; h \in \xi \qquad \text{(Eq. A.6)}$$

$$\mathfrak{I}_{1,i} = \{f_i(x) = (1-\varepsilon_1)f_{1,i}(x) + \ \varepsilon_1 h(x) \ \ h \in \xi \qquad \text{(Eq. A.7)}$$

where ξ is the class of all one-dimensional density functions within $\mathfrak{R}$, then the least favorable density pair is

$$f_{o,i}^{*}(x) = \begin{cases} (1-\varepsilon_o)f_{o,i}(x) & x\colon \dfrac{f_{1,i}(x)}{f_{o,i}(x)} < c_{o,i} \\ c_o^{-1}(1-\varepsilon_o)f_{1,i}(x) & x\colon \dfrac{f_{1,i}(x)}{f_{o,i}(x)} \geq c_{o,i} \end{cases} \qquad \text{(Eq. A.8)}$$

$$f_{1,i}^{*}(x) = \begin{cases} (1-\varepsilon_1)f_{1,i}(x) & x\colon \dfrac{f_{1,i}(x)}{f_{o,i}(x)} > c_{1,i} \\ c_1(1-\varepsilon_1)f_{o,i}(x) & x\colon \dfrac{f_{1,i}(x)}{f_{o,i}(x)} \leq c_{o,i} \end{cases} \qquad \text{(Eq. A.9)}$$

where $c_{0,i}, c_{1,i}$ are such that

$$\Pr\left\{\left(\frac{f_{1,i}(X)}{f_{o,i}(X)}\right) < c_{o,i}\middle| f_{o,i}\right\} + \mathrm{c}_{\mathrm{o},i}^{-1}\Pr\left\{\left(\frac{f_{1,i}(X)}{f_{o,i}(X)}\right) \geq c_{0,i}\middle| f_{1,i}\right\} = (1-\varepsilon_0)^{-1} \qquad \text{(Eq. A.10)}$$

$$\Pr\left\{\left(\frac{f_{1,i}(X)}{f_{o,i}(X)}\right) > c_{1,i}\middle| f_{1,i}\right\} + \mathrm{c}_{1,\mathrm{i}}\Pr\left\{\left(\frac{f_{1,i}(X)}{f_{o,i}(X)}\right) \leq c_{1,i}\middle| f_{0,i}\right\} = (1-\varepsilon_1)^{-1} \qquad \text{(Eq. A.11)}$$

If we consider a robust detection rule over $\mathfrak{I}_1^n$ and $\mathfrak{I}_o^n$, where $\mathfrak{I}_{1,i}^n$ and $\mathfrak{I}_{o,i}^n$ are the Huber's classes defined by (Eq. A.6) and (Eq. A.7) and where $\mathfrak{I}_1^n$ and $\mathfrak{I}_o^n$ contain only memoryless processes, then, due to Lemma A.1 and (Eq. A.8) and (Eq. A.9), our robust detection rule is

$$\delta^{*}(x^n) = \begin{cases} 1 & x^n\colon T^{*}(x^n) > \lambda^{*} \\ r^{*} & x^n\colon T^{*}(x^n) = \lambda^{*} \\ 0 & x^n\colon T^{*}(x^n) < \lambda^{*} \end{cases} \qquad \text{(Eq. A.12)}$$

where r^*, λ^* are selected to satisfy the false alarm constraint, and where

$$T^*(x^n) = n^{-1} \sum_{i=1}^{n} \gamma_i^*(x_i) \qquad \text{(Eq. A.13)}$$

$$\gamma_i^*(x_i) = \log \frac{(f_{1,i}^*(x)}{f_{0,i}^*(x)}$$

$$= \begin{cases} \log\left[c_{1,i}(1-\varepsilon_1)(1-\varepsilon_o)^{-1}\right] & x: f_{1,i}(x)/f_{o,i}(x) \leq c_{1,i} \\ \log\left[(1-\varepsilon_1)(1-\varepsilon_o)^{-1} f_{1,i}(x)/f_{o,i}(x)\right] & x: c_{1,i} < f_{1,i}(x)/f_{o,i}(x) < c_{0,i} \\ \log\left[c_{o,i}(1-\varepsilon_1)(1-\varepsilon_o)^{-1}\right] & x: f_{1,i}(x)/f_{o,i}(x) \geq c_{0,i} \end{cases} \qquad \text{(Eq. A.14)}$$